

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank) 2. REPORT DATE 97 3. REPORT TYPE AND DATES COVERED
24 December Final 1 November 1996-30 October 1997

4. TITLE AND SUBTITLE
Final Report:
Twenty-fourth Annual Conference on the Physics and
Chemistry of Semiconductor Interfaces

5. FUNDING NUMBERS

G: N00014-97-1-0033

6. AUTHOR(S)

C. R. Schulte

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Institute for Postdoctoral Studies
P. O. Box 36
Scottsdale, AZ 85252-0036

8. PERFORMING ORGANIZATION
REPORT NUMBER

FI-1997

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
ATTN: ONR 251:JWK
Ballston Tower One
800 North Quincy Street
Arlington, VA 22217-5660 CFDA No. 12.300

10. SPONSORING/MONITORING
AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Office of Naval Research position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The Twenty-fourth Annual Conference on the Physics and Chemistry of Semiconductor Interfaces was held in Research Triangle Park, NC and the enclosed report contains abstracts of papers presented.

14. SUBJECT TERMS

Semiconductor Interfaces

15. NUMBER OF PAGES

276

16. PRICE CODE

17. SECURITY CLASSIFICATION
OF REPORT

Unclassified

18. SECURITY CLASSIFICATION
OF THIS PAGE

Unclassified

19. SECURITY CLASSIFICATION
OF ABSTRACT

Unclassified

20. LIMITATION OF ABSTRACT

UL

Institute for Postdoctoral Studies

C. R. Schulte
President
Institute for Postdoctoral Studies
P. O. Box 36
Scottsdale, Arizona 85252-0036
Phone: (602) 423-8540
FAX: (602) 423-5183

24 December 1997

Defense Technical Information Center
8725 John J. Kingman Road
STE 0944
Ft. Belvoir, Virginia 22060-6218

RE: Final Report of N00014-97-1-0033

Dear Sir or Madam:

Enclosed is a copy of the Final Report for Grant Number N00014-97-1-0033, which includes pages 1015-1285 of Volume B15 (4) of the Journal of Vacuum Science and Technology B.

Thank you for your support.

Sincerely,



C. R. Schulte
President

CRS

Enclosures

OFFICE OF NAVAL RESEARCH

FINAL REPORT

for

Grant No. N00014-97-1-0033

PR Number 97PR00764-00

**Twenty-fourth Annual Conference of the Physics and Chemistry of
Semiconductor Interfaces**

Under the direction of
C. R. Schulte
Institute for Postdoctoral Studies
P. O. Box 36
Scottsdale, Arizona 85252-0036

Reproduction in whole, or in part, is permitted for any purpose of the United States Government.

* This document has been approved for public release and sale; its distribution is unlimited.

report.onr97pcsi.

DATA QUALITY INSPECTED 4

**Twenty-fourth Annual Conference of the Physics and Chemistry of
Semiconductor Interfaces**

Under the direction of
C. R. Schulte
Institute for Postdoctoral Studies
P. O. Box 36
Scottsdale, Arizona 85252-0036

Conference

The Twenty-fourth Annual Conference on the Physics and Chemistry of Semiconductor Interfaces was held in Research Triangle Park, North Carolina in January of 1997, and was attended by over 100 workers in the field of semiconductor interfaces, with researchers from all over the world.

The conference featured extensive discussion sessions and a workshop format, and was judged a success by its attendees — largely because it facilitated informal discussions among the attendees, and especially involving young attendees (*e.g.*, graduate students) interacting with senior scientists.

Publications

The Proceedings of the conference are enclosed, and feature the papers presented which have been published in the Journal of Vacuum Science and Technology.

Expenditures of funds

The following represents the expenditure of Department of Defense funds for this conference:

EXPENDITURES OF DEPARTMENT OF DEFENSE FUNDS: ONR**A. Travel**

Domestic travel by Committee: \$6,058.25

Domestic travel and conference fees for speakers/discussion leaders: 25.00

TOTAL TRAVEL 6,083.25**B. Secretarial Services** 4,990.00**C. Publications** 230.00**D. Graphics, Printing, Postage, Communications** 1,500.00**E. Audio-visual equipment and related facilities** 1,000.00**F. All other conference facilities and services** 1,196.75**G. Accounting, overhead, and legal services (a flat fee)** 1,000.00

H. GRAND TOTAL \$16,000.00

FINANCIAL STATUS REPORT (Long Form)

(Follow instructions on the back)

1. Federal Agency and Organizational Element to Which Report is Submitted Office of Naval Research San Diego, CA 92121-3019		2. Federal Grant or Other Identifying Number Assigned By Federal Agency N00014-97-1-0033		OMB Approval No. 0348-0039	Page 1	of 1 pages
3. Recipient Organization (Name and complete address, including ZIP code) Institute for Postdoctoral Studies P. O. Box 36 Scottsdale, AZ 85252-0036						
4. Employer Identification Number 86-0760869		5. Recipient Account Number or Identifying Number PCSI24		6. Final Report <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No		7. Basis <input checked="" type="checkbox"/> Cash <input type="checkbox"/> Accrual
8. Funding/Grant Period (See Instructions) From: (Month, Day, Year) November 01 1996		To: (Month, Day, Year) October 30, 1997		9. Period Covered by this Report From: (Month, Day, Year) Nov. 01, 1996		To: (Month, Day, Year) Oct. 30, 1997
10. Transactions:						
		Previously Reported	This Period	Cumulative		
a. Total outlays		14400.00	1600.00	16000.00		
b. Refunds, rebates, etc.		0	0	0		
c. Program income used in accordance with the deduction alternative						
d. Net outlays (Line a, less the sum of lines b and c)		14400.00	1600.00	16000.00		
Recipient's share of net outlays, consisting of:						
e. Third party (in-kind) contributions		0	0	0		
f. Other Federal awards authorized to be used to match this award		0	0	0		
g. Program income used in accordance with the matching or cost sharing alternative		0	0	0		
h. All other recipient outlays not shown on lines e, f or g		0	0	0		
i. Total recipient share of net outlays (Sum of lines e, f, g and h)		0	0	0		
j. Federal share of net outlays (line d less line i)				16000.00		
k. Total unliquidated obligations					0	
l. Recipient's share of unliquidated obligations					0	
m. Federal share of unliquidated obligations					0	
n. Total federal share (sum of lines j and m)					16000.00	
o. Total federal funds authorized for this funding period					16000.00	
p. Unobligated balance of federal funds (Line o minus line n)					0	
Program income, consisting of:					0	
q. Disbursed program income shown on lines c and/or g above					0	
r. Disbursed program income using the addition alternative					0	
s. Undisbursed program income					0	
t. Total program income realized (Sum of lines q, r and s)					0	
11. Indirect Expense						
a. Type of Rate (Place "X" in appropriate box) <input type="checkbox"/> Provisional <input checked="" type="checkbox"/> Predetermined <input type="checkbox"/> Final <input type="checkbox"/> Fixed						
b. Rate \$1000.00 flat		c. Base \$16000.00		d. Total Amount \$1000.00		e. Federal Share \$1000.00
12. Remarks: Attach any explanations deemed necessary or information required by Federal sponsoring agency in compliance with governing legislation.						
13. Certification: I certify to the best of my knowledge and belief that this report is correct and complete and that all outlays and unliquidated obligations are for the purposes set forth in the award documents.						
Typed or Printed Name and Title C. R. Schulte, Principal Investigator				Telephone (Area code, number and extension) 602-423-8540		
Signature of Authorized Certifying Official <i>C R Schulte</i>				Date Report Submitted 24 December 1997		

Journal of Vacuum Science & Technology B

JVST B

Second Series
Volume 15, Number 4
July/August 1997

Microelectronics and Nanometer Structures

Processing, Measurement, and Phenomena

**Papers from the 24th Annual Conference
on the Physics and Chemistry
of Semiconductor Interfaces**

**Papers from the Fourth International Conference
on Nanometer-Scale Science and Technology**

19971230 005



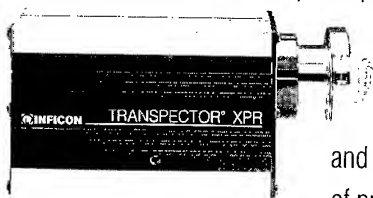
An official journal of the American Vacuum Society
Published for the Society by the American Institute of Physics



Relax. PVD pressures can't torture our gas analyzer.

Because of its superior measurement range, Transpector® XPR doesn't need complex pumping systems.

Transpector XPR is the first quadrupole-based gas analyzer to operate at PVD process and base pressures, without the cost, complexity and maintenance requirements of pressure-conversion systems.



A remarkable miniaturized quadrupole, reduced 10 times in size, allows Transpector XPR to operate within a pressure range of 20 mtorr to ultrahigh vacuum. So it can monitor PVD processes for contamination control.

This incredibly simple design makes Transpector XPR compact and affordable enough to be installed on each chamber of your PVD tool. Add TranspectorWare™ process gas characterization software and you'll have real-time security windows into all stages of your process, for increased uptime and improved wafer yield.

Get a sharper view of your process, without the torture of complex pumping systems. For more information on what Transpector XPR can do for you, call today.



LEYBOLD INFICON

The Instrumental Difference™

United States

Phone: (315) 434-1100
Fax: (315) 437-3803

Germany

Phone: (+49) 221 347-0
Fax: (+49) 221 347-1250

France

Phone: (+33) 01 69 076400
Fax: (+33) 01 69 075738

United Kingdom

Phone: (+44) 0181 9717000
Fax: (+44) 0181 9717005

Netherlands

Phone: (+31) 03480 77411
Fax: (+31) 03480 20489

Japan

Phone: (+81) 045 471 3311
Fax: (+81) 045 471 3324

Korea

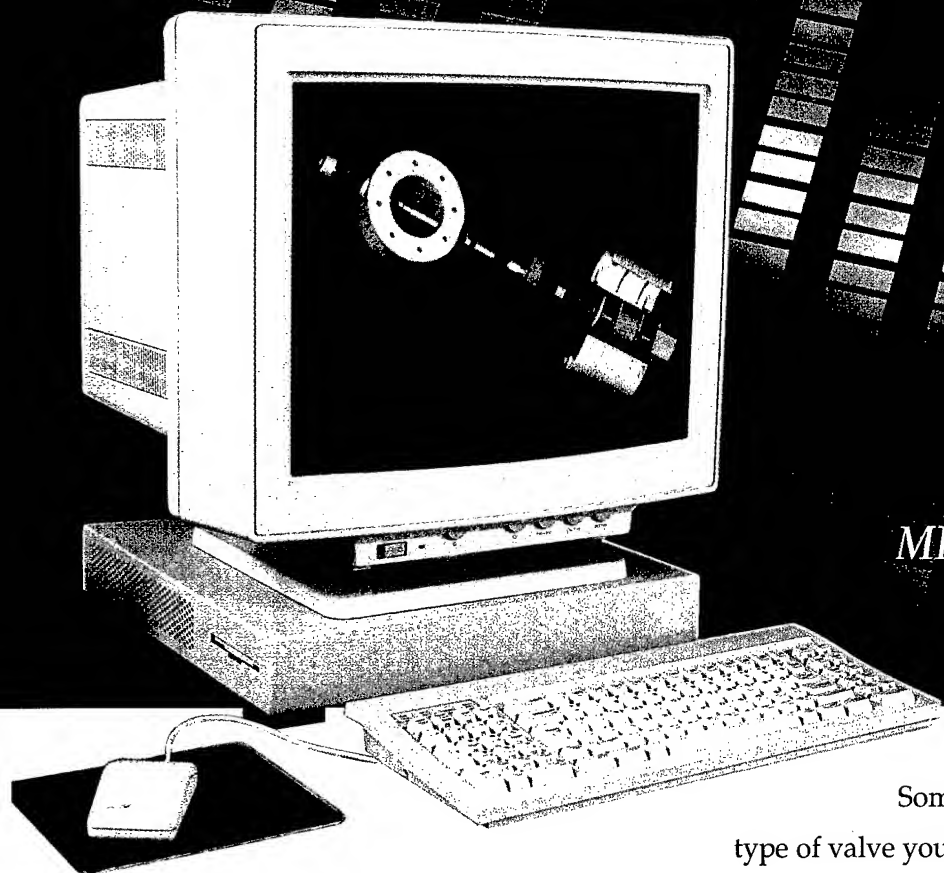
Phone: (+82) 02 270 1315
Fax: (+82) 02 271 1568

Taiwan

Phone: (+886) 03 5425670
Fax: (+886) 03 5421233

Visit our website for other sales offices worldwide. <http://www.leyboldinficon.com>

E-mail: reachus@inficon.com



*MKS Around the World ...
... Around the Process.*

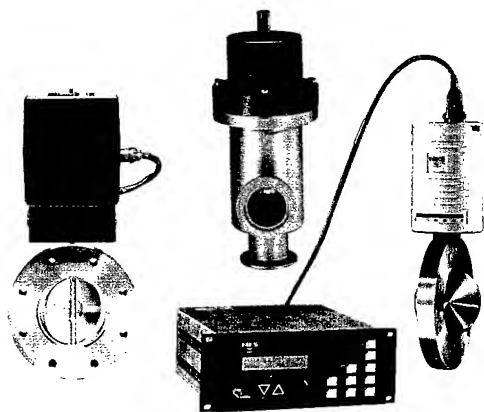
*"Choosing the
right valve is
easy, once you
choose MKS!"*

Sometimes we think we offer every
type of valve you'll need to control your process,

but we don't. That's why our design services group stays so
busy. Our design engineers work with OEM's every day who
have "new" applications. These customers turn to MKS for the
industry's widest selection of standard features and custom-
engineered options, including:

- smart throttle valves with analog, serial, or network (Sensor Bus) control
- butterfly valves that both seal and control
- high torque direct drives for non-slip control
- heaters for operation to 250°C
- advanced algorithm self-tuning valve controllers

MKS offers the best technology platform to handle anyone's
standard and custom pressure control requirements. Choose
the right valve, choose MKS. Call us today for detailed product
literature, applications assistance or design services at
(800) 227-8766/(508) 975-2350.



MKS Instruments, Inc. Six Shattuck Road, Andover, MA 01810 Tel: (800)227-8766/(508)975-2350 Fax: (508)975-0093
e-mail: mks@mksinst.com www: <http://www.mksinst.com>

1997 MKS Instruments, Inc. ©. All rights reserved.

Journal of Vacuum Science & Technology B Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena

JVST B

ISSN: 0734-211X
CODEN: JVTBD9

American Vacuum Society

Officers

Gary E. McGuire, *President*
Microelectronics Center
of North Carolina

Jerry M. Woodall, *President-Elect*
Purdue University

William D. Sproul, *Immediate
Past-President*
Sputtered Films

Joseph E. Greene, *Secretary*
University of Illinois

N. Rey Whetten, *Treasurer*
American Vacuum Society

Directors

Dawn A. Bonnell
University of Pennsylvania

Richard A. Hoffman
ALCOA

Rudolf Ludeke
IBM T. J. Watson Research Center

Christie Marrian
Naval Research Laboratory

Dale C. McIntyre
QM Technologies

Ellen B. Stechel
Sandia National Laboratories

JVST

G. Lucovsky, *Editor-in-Chief*
North Carolina State University

**AVS Membership
Information may be
obtained from**

Angela Mulligan
*AVS Membership
and Scholarship Coordinator*
120 Wall Street
32nd Floor
New York, NY 10005
(212) 248-0200

Editor: G. E. McGuire,
Microelectronics Center of North Carolina,
Research Triangle Park, NC 27709. Telephone: (919) 248-1910

Supervisor Editorial Office: Rebecca York,
Editorial Assistant: Mary P. Hipsher,
Journal of Vacuum Science and Technology, Microelectronics Center of North Carolina,
3021 Cornwallis Rd., Caller Box 13994, Research Triangle Park, NC 27709-3994.
Telephone: (919) 248-1861 and (919) 248-1884; FAX: (919) 248-1857; E-mail: jvst@mcnc.org

Associate Editors:

Eric Kay, Critical Reviews and Special Issues
C. R. Abernathy (98) University of Florida
P. C. Fazan (97) Micron Technology, Inc.
J. Givens (99) Micron Technology, Inc.

B. Gnade (97) Texas Instruments, Inc.
J. L. Shaw (98) Naval Research Lab.
D. Temple (99) MCNC

JVST Publication Committee:

J. William Rogers, Jr., Chair, University of
Washington

John E. Crowell, UC San Diego

Joseph D. Geller, Geller Microanalytical Lab.
Pete Sheldon, NREL
Dorota Temple, MCNC

JVST Editorial Board:

Linda M. Cecchi (97) Sandia National Lab.
S. A. Chambers (99) Pacific NW National Lab.
R. J. Colton (98) Naval Research Lab.
T. M. Mayer (97) Sandia National Lab.

B. S. Meyerson (98) IBM Watson Research Ctr.
Buddy D. Ratner (97) University of Washington
Gary W. Rubloff (97) University of Maryland
D. Ruzic (99) University of Illinois
Bruce D. Sartwell (97) Naval Research Lab.

JVST Editorial Staff at AIP: Editorial Supervisor: Deborah McHone; Journal Coordinator: Margaret
Reilly; Chief Production Editor: Mary Ellen Mornile

The Journal of Vacuum Science and Technology B is published six times annually (Jan/Feb, Mar/Apr, May/June, Jul/Aug, Sep/Oct, Nov/Dec) by the American Institute of Physics (AIP) for and under the editorial management of the American Vacuum Society (AVS). It is an official publication of the AVS and is received by all members of the Society. It is devoted to reports of original research, review articles, and Critical Review articles. The *JVST B* has been established to provide a vehicle for the publication of research dealing with microelectronics and nanometer structures. The emphasis will be on processing, measurement, and phenomena, and will include vacuum processing, plasma processing, materials and structural characterization, microlithography, and the physics and chemistry of submicron and nanometer structures and devices. This journal will publish the proceedings of conferences and symposia that are sponsored by the AVS and its divisions.

Submit Manuscripts to the Editorial Office of the *Journal of Vacuum Science and Technology*, Microelectronics Center of North Carolina, 3021 Cornwallis Rd., Caller Box 13994, Research Triangle Park, North Carolina 27709-3994. Manuscripts of papers presented at AVS-sponsored conferences and symposia and being submitted to *JVST B* should be sent to the Guest Editor appointed for that particular conference. Before preparing a manuscript, authors should read "Information for Contributors," printed in the first issue of each volume of the journal. Submission of a manuscript is a representation that the manuscript has not been published previously nor currently submitted for publication elsewhere. Upon receipt of a manuscript the Editor will send the author a Transfer of Copyright Agreement form. This must be completed by the author and returned *only* to the Editorial Office prior to publication of an accepted paper in the *Journal of Vacuum Science and Technology B*. This written transfer of copyright, which previously was assumed to be implicit in the act of submitting a manuscript, is necessary under the 1978 copyright law in order for the AVS and AIP to continue disseminating research results as widely as possible. Further information may be obtained from AIP.

Publication Charge: To support the cost of wide dissemination of research results through the publication of journal pages and production of a database of articles, the author's institution is requested to pay a *page charge* of \$95 per page (with a one-page minimum). The charge (if honored) entitles the author to 100 free reprints. For Errata the minimum page charge is \$10, with no free reprints.

Physics Auxiliary Publication Service (PAPS) and/or Electronic PAPS (E-PAPS): For a nominal fee, authors may submit material that is part of and supplemental to a paper, but is too long to be included in the journal. PAPS deposits may be on paper or in electronic media, and can include text, data in graphic or numeric form, computer programs, etc. Retrieval instructions are footnoted in the related published paper. Direct requests to the Editor.

Advertising Rates will be supplied on request from AIP's Advertising Division, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999. Telephone: (516) 576-2440. Fax: (516) 576-2481. E-mail: advtsg@aip.org. All insertion orders and advertising material should be sent to that division.

Copying: Single copies of individual articles may be made for private use or research. Authorization is given (as indicated by the Item Fee Code for this publication) to copy articles beyond the use permitted by Sections 107 and 108 of the U.S. Copyright Law, provided the copying fee of \$10 per copy per article is paid to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, USA. Persons desiring to photocopy materials for classroom use should contact the CCC Academic Permissions Service. The Item Fee Code for this publication is 0734-211X/97 \$10.00.

Authorization does not extend to systematic or multiple reproduction, to copying for promotional purposes, to electronic storage or distribution, or to republication in any form. In all such cases, specific written permission from AIP must be obtained.

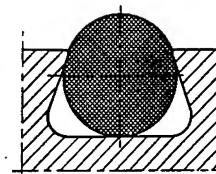
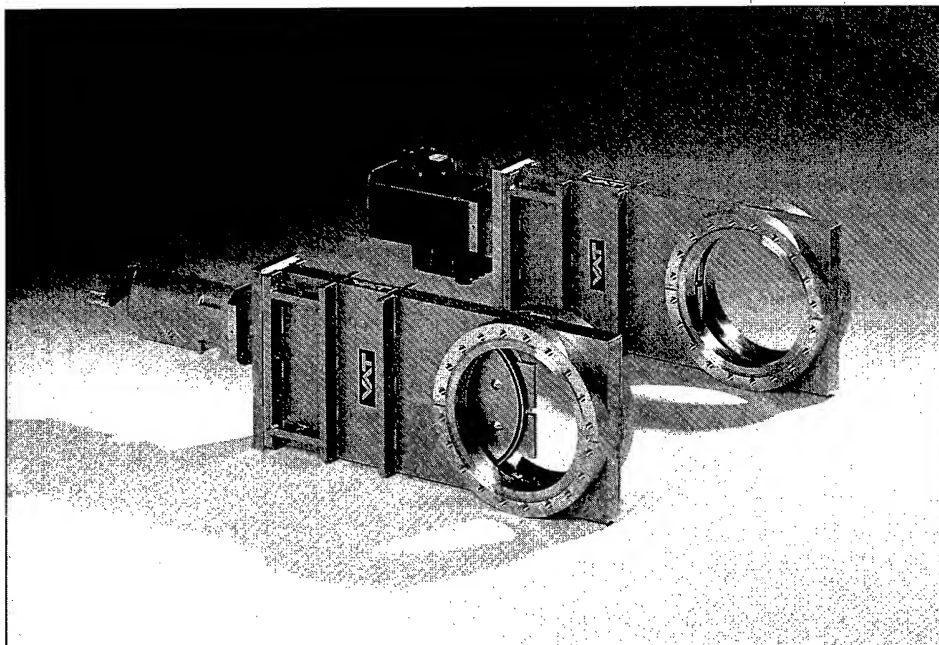
Permission for Other Use: Permission is granted to quote from the journal with the customary acknowledgment of the source. To reprint a figure, table, or other excerpt requires the consent of one of the authors and notification to AIP.

Requests for Permission: Address requests to AIP Office of Rights and Permissions, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999; Telephone: (516) 576-2268; Fax: (516) 576-2327; Internet: rights@aip.org.

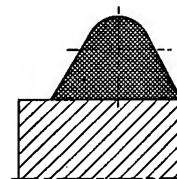
Copyright © 1997 American Vacuum Society. All rights reserved.

®

Innovation in
UHV Gate Valves
Series 10 with
"Vulcanized" Gate Seal



Conventional
O-Ring Gate
Seal



VAT Vulcanized
UHV Gate
Seal

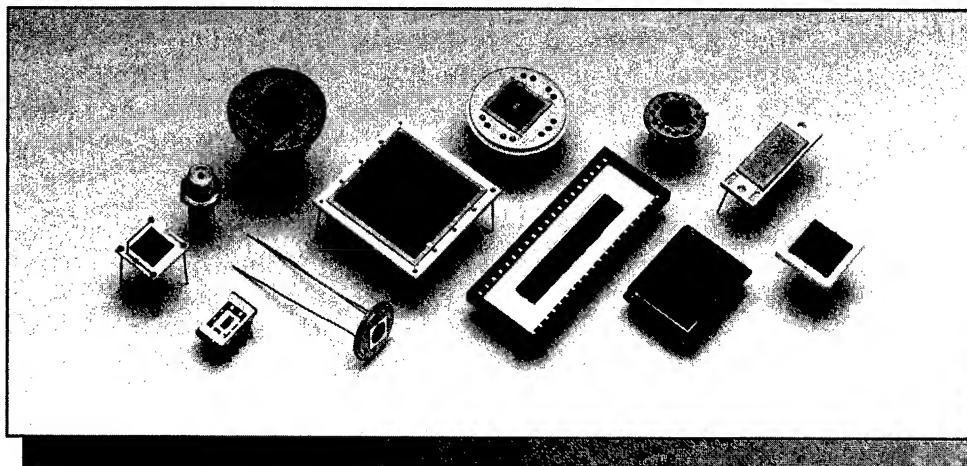
- One Sealing Line Only
- No Trapped Volumes
- Higher Lifetime

Request your new Catalog 2000 today

VAT, Inc., 500 West Cummings Park, Woburn, MA 01801
800-935-1446 Fax 617-935-3940
e mail: usa@vatvalve.com <http://www.vatvalve.com>

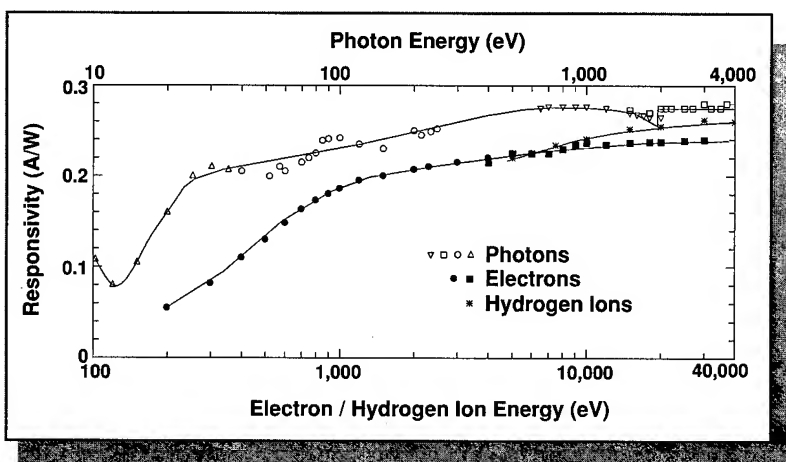
ABSOLUTE

XUV SILICON PHOTODIODES



Features

- ◆ 100% Internal Quantum Efficiency
- ◆ 8 Decades of Linearity
- ◆ No Change in QE with 100 Mrad (Si) 124 eV Photons
- ◆ May be Operated Without Bias
- ◆ Cryogenically and UHV Compatible
- ◆ With Integrated Bandwidth Limiting Filters



TYPICAL RESPONSIVITY OF AXUV PHOTODIODES TO PHOTONS, ELECTRONS AND HYDROGEN IONS

INTERNATIONAL RADIATION DETECTORS INC.

2545 West 237th Street, Unit I ♦ Torrance, California 90505-5229
(310) 534-3661 ♦ FAX (310) 534-3665 ♦ email: ird@kaiwan.com



DEVELOPED IN COLLABORATION WITH NIST, NIH, LLNL, LANL, NCAR

Journal of Vacuum Science & Technology B

Microelectronics and Nanometer Structures

Processing, Measurement, and Phenomena

JVST B

Second Series
Volume 15, Number 4
Jul/Aug 1997

Review Article

Microstructure control in semiconductor metallization

J. M. E. Harper and K. P. Rodbell 763

Regular Articles

Dual unit scanning tunneling microscope-atomic force microscope for length measurement based on reference scales

Haijun Zhang, Feng Huang, and Toshiro Higuchi 780

(Continued)

Journal of Vacuum Science & Technology B (ISSN: 0734-211X) is published six times annually (Jan/Feb, Mar/Apr, May/Jun, Jul/Aug, Sep/Oct, Nov/Dec) for the American Vacuum Society by the American Institute of Physics, 500 Sunnyside Blvd., Woodbury, NY 11797-2999. Membership in the American Vacuum Society includes \$17.50 from membership dues to be applied towards a subscription to *Journal of Vacuum Science & Technology B*.

Known office of publication is Thorofare, NJ 08086. Periodicals postage paid at Thorofare, NJ 08086, and at additional mailing offices. POSTMASTER: Send address changes to *Journal of Vacuum Science & Technology B*, SLACK Inc., 6900 Grove Road, Thorofare, NJ 08086.

Subscription Prices (1997)

	U.S.A. and Poss.	Can., Mex., Central & S. America & Caribbean	Europe, Asia, Africa & Oceania*
JVST B [†]	\$670	\$695	\$714
JVST B [‡]	\$770	\$795	\$814
JVST A and B [†]	\$770	\$820	\$858
JVST A and B [¶]	\$770	\$795	\$795
JVST A and B [‡]	\$870	\$920	\$958

[†]Paper only.

[‡]Paper and CD-ROM.

[¶]CD-ROM only.

*Nonmember subscriptions include air-freight service.

Back-number Prices: 1997 single copy: \$125. Prior to 1997 single copies: \$105.

Subscriptions, renewals, and address changes should be addressed to *Subscription Fulfillment Division, SLACK Inc., 6900 Grove Road, Thorofare, NJ 08086*. Allow at least six weeks advance notice. For address changes please send both old and new addresses, and, if possible, include a label from the plastic mailing wrapper of a recent issue. For your convenience a **change-of-address form is included in every issue of *Physics Today*; please use it**. AIP will honor a request for a missing journal issue only within six months of that issue's actual date of publication. Requests received for missing issues beyond six months of the actual publication date will not be honored. The issue may, however, be purchased at the single-copy price. (Subscription Fulfillment offices are located at SLACK Inc., 6900 Grove Road, Thorofare, NJ 08086.)

Single-copy orders (current and back issues) should be addressed to American Institute of Physics, Circulation and Fulfillment Division, 500 Sunnyside Blvd., Woodbury, NY 11797-2999.

Reprints: Reprints can be ordered with or without covers only in multiples of 50 (with a minimum of 100 in each category) from AIP, Circulation & Fulfillment/Reprints, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999; Fax: 516-349-9704; Telephone: 800-344-6909 (in U.S. and Canada), or 516-576-2234.

Document Delivery: Copies of articles can be ordered for \$15 per copy from the AIP/Member Society document delivery service, "Articles in Physics," 75 Varick Street, New York, NY 10013; Fax: 212-301-4060; Telephone: 800-480-PHYS (800-480-7497) (in U.S. and Canada), or 212-301-4000; E-mail: articles@aip.org; URL: <http://www.aip.org/articles.html>.

Copying fees: The code that appears on the first page of articles in this journal gives the fee for each copy of the article made beyond the free copying permitted by AIP. (See statement under "Copyright" elsewhere in this journal.) If no code appears, no fee applies. The fee for pre-1978 articles is \$0.25 per copy. With the exception of copying for advertising and promotional purposes, the express permission of AIP is not required provided the fee is paid through the *Copyright Clearance Center, Inc. (CCC), 222 Rosewood Dr., Danvers, MA 01923*. Contact the CCC for information on how to report copying and remit payment.

Microform: *Journal of Vacuum Science and Technology B* is available on microfiche issued at the same frequency as the printed journal and annually on microfilm. A microform catalog is available from AIP, Circulation & Fulfillment/Single, Copy Sales, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999; Fax: 516-349-9704; Telephone: 800-344-6908 (in U.S. and Canada), or 516-576-2277.

Online Availability: Abstracts of journal articles published by AIP and Member Societies (and several other physics publishers) are available in the SPIN database via the AIP online service PINET. Also available on PINET are *Advance Abstracts*, a current awareness service, other physics information resources, as well as internet services (Internet: elecprod@aip.org).

Magnetic nanostructures fabricated by scanning tunneling microscope-assisted chemical vapor deposition	
Woei Wu Pai, Jiandi Zhang, John F. Wendelken, and R. J. Warmack.....	785
Force modulation atomic force microscopy recording for ultrahigh density recording	
S. Hosaka, H. Koyanagi, A. Kikukawa, M. Miyamoto, K. Nakamura, and K. Etoh.....	788
Nanofabrication of electrodes with sub-5 nm spacing for transport experiments on single molecules and metal clusters	
A. Bezryadin and C. Dekker.....	793
Deconvolution of tip affected atomic force microscope images and comparison to Rutherford backscattering spectrometry	
M. F. Tabet and F. K. Urban III.....	800
Effect of substrate temperature and annealing on the structural properties of ZnO ultrafine particle films	
Zhao Dachun, Qu Zhongkai, Pan Xiaoren, Dai Muji, and Sun Minggen.....	805
Ion implanted nanostructures on Ge(111) surfaces observed by atomic force microscopy	
Y. J. Chen, I. H. Wilson, W. Y. Cheung, J. B. Xu, and S. P. Wong.....	809
Nanometer table-top proximity x-ray lithography with liquid-target laser-plasma source	
L. Malmqvist, A. L. Bogdanov, L. Montelius, and H. M. Hertz.....	814
Low-energy focused-ion-beam exposure characteristics of an amorphous Se₇₅Ge₂₅ resist	
Hyun-Yong Lee and Hong-Bay Chung.....	818
Novel effects in inorganic As₅₀Se₅₀ photoresists and their application in micro-optics	
V. Lyubin, M. Klebanov, I. Bar, S. Rosenwaks, N. P. Eisenberg, and M. Manevich.....	823
Imaging characteristics of poly(methyl methacrylate) at vacuum ultraviolet wavelengths	
I. E. Ferincz, Cs. Tóth, and J. F. Young.....	828
Design optimization for two lens focused ion beam columns	
Li Wang.....	833
Field emission from amorphous diamond coated Mo tip emitters by pulsed laser deposition	
M. Q. Ding, A. F. Myers, W. B. Choi, R. D. Vispute, S. M. Camphausen, J. Narayan, J. J. Cuomo, J. J. Hren, and J. Bruley.....	840
Oxygen implantation induced interdiffusion in AlGaAs/GaAs quantum well structures	
Padraig J. Hughes, Bernard L. Weiss, Spirit Tlali, and Howard E. Jackson.....	845
Observation of hot electron relaxation in GaAs/AlGaAs multiple quantum wells by excitation spectroscopy	
H. Z. Wu, J. H. Liu, G. O. Dong, J. Z. Wu, Z. Z. Ye, and X. B. Jiang.....	849
Base metallization stability in InP/InGaAs heterojunction bipolar transistors and its influence on leakage currents	
D. Caffin, C. Besombes, J. F. Bresse, P. Legay, G. Le Roux, G. Patriarche, and P. Launay.....	854
Structural and optical characterizations of single three-dimensionally confined GaAs/AlAs structures grown on patterned GaAs (001) substrates	
Jun-ichi Kasai, Sige-hisa Tanaka, Katsuhiko Higuchi, and Yoshifumi Katayama.....	862
Electrical transport properties of silicon delta-doped Al_{0.30}Ga_{0.70}As samples showing suppression of the DX center features	
J. A. Corrêa F., A. G. de Oliveira, M. I. N. da Silva, G. M. Ribeiro, and J. F. Sampaio.....	870
X-ray photoelectron spectroscopy study of GaAs(110) cleaved in alcoholic sulfide solutions	
Vasily N. Bessolov, Elena V. Konenkova, and Mikhail V. Lebedev.....	876
Effects of water vapor and chlorine on the epitaxial growth of Si_{1-x}Ge_x films by chemical vapor deposition: Thermodynamic analysis	
I.-M. Lee, A. Jansons, and C. G. Takoudis.....	880
Integration issues for 850 nm optical modulators on Si electronics by direct epitaxy	
J. E. Cunningham and W. Y. Jan.....	886
Interfacial reaction behavior of Pt, Pd, and Ni on ZnSe	
K. J. Duxstad, E. E. Haller, K. M. Yu, E. D. Bourret, X. W. Lin, S. Ruvimov, Z. Liliental-Weber, and J. Washburn.....	891
Ultrathin cobalt silicide layers formed by rapid thermal processing of metal on amorphous silicon	
A. H. M. Kamal, M. J. Rack, M. N. Kozicki, D. K. Ferry, J. Lützen, and J. A. Hallmark.....	899
Electrical characterization of iridium Schottky contacts to silicon: Early stages of silicidation	
F. J. Jiménez-Leube, M. Clement, J. Sanz Maudes, and T. Rodríguez.....	903

(Continued)

Mass Spectrometers for Gas Analysis



process
monitoring



research



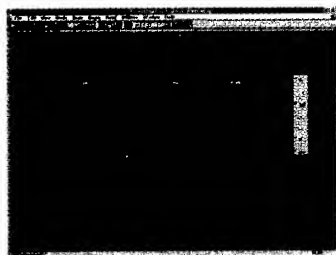
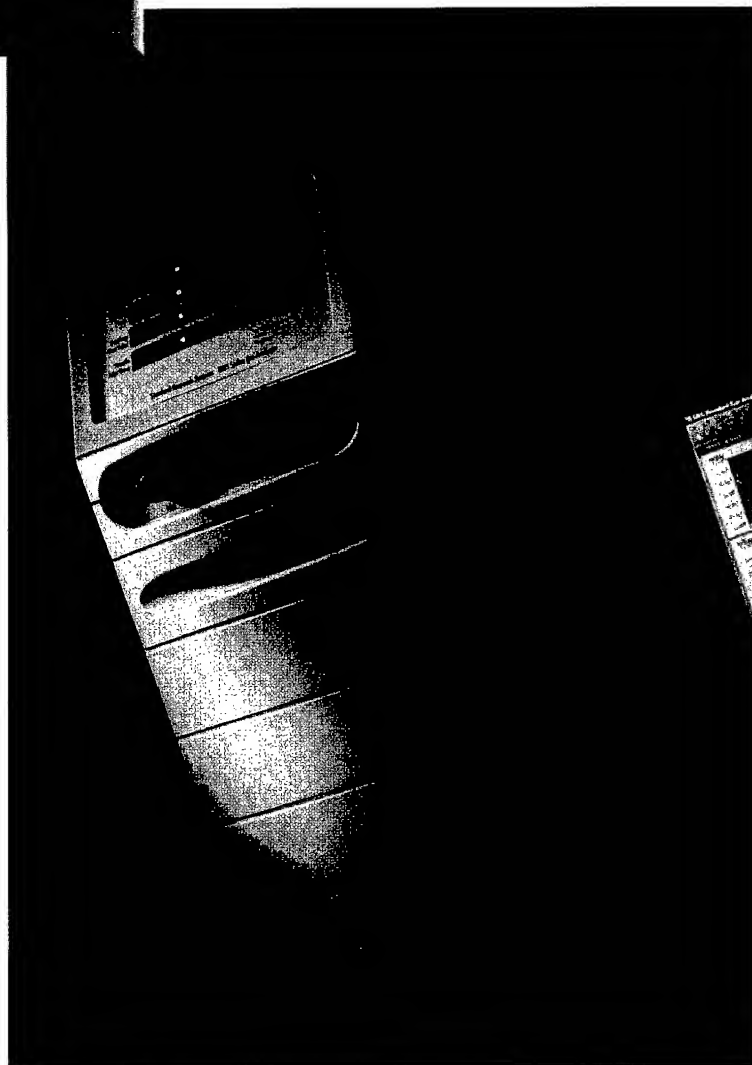
leak
detection

- Response time of less than 1 second
- Compact, transportable design
- Large dynamic range allows detection to ppm levels
- Real-time Windows® analysis software
- Simple to setup and use

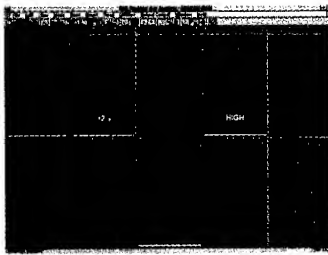
The QMS Series Gas Analyzers give you the power of on-line gas analysis and monitoring for a wide range of applications. These self-contained automated systems house a modern mass spectrometer for the analysis of light gases. You can detect leaks, measure gas composition and troubleshoot process lines.

The QMS system continuously samples gases, as opposed to batch sampling, which is utilized by gas chromatographs. This allows the QMS to detect gases at high speed – response time is less than 1 second. The inlet pressure can be configured to measure from vacuum to above atmospheric pressure.

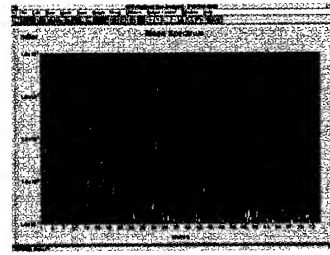
A Windows® based software package displays data in real-time and is included with each instrument. Offered as 100, 200 or 300 amu systems, one of our QMS Gas Analyzers will suit your application. Call SRS for a full brochure on the QMS or any of our other gas analysis instruments.



Pressure vs. time mode presents a strip chart display of partial pressures for selected gases and provides a complete history of data.



Annunciator mode monitors up to 10 gases. User set limits allow Go/No-Go testing in conjunction with audible alarms and/or an optional relay board.



Analog mode presents the entire spectrum and displays the results as partial pressure vs. mass number. Gas monitoring can be continuous or timed.



Stanford Research Systems

1290-D Reamwood Avenue • Sunnyvale, CA 94089

Telephone: (408)744-9040 • FAX: (408)744-9049 • Email: info@srsys.com • WWW: <http://www.srsys.com>

Linewidth dependence of stress relaxation and microstructural change in passivated Al(Cu) lines I.-S. Yeo, P. S. Ho, S. G. H. Anderson, and H. Kawasaki	908
Probeless voltage contrast using a focused ion beam for opens and shorts defect isolation of ultralarge scale integration technologies Kenneth J. Giewont, Douglas B. Hunt, and Klaus M. Hummler.....	916
Surface roughness of nitrided (0001) Al₂O₃ and AlN epilayers grown on (0001) Al₂O₃ by reactive molecular beam epitaxy Wook Kim, M. Yeadon, A. E. Botchkarev, S. N. Mohammad, J. M. Gibson, and H. Morkoç.....	921
Influences of the [(Ba,Sr)TiO₃]-modified RuO₂ interface on the dielectric constant and current-voltage characteristics Min-Seok Jeon and Duck-Kyun Choi.....	928
Equipment simulation of SiGe heteroepitaxy: Model validation by <i>ab initio</i> calculations of surface diffusion processes M. Hierlemann, C. Werner, and A. Spitzer.....	935
Physical and electrical properties in metal-oxide-Si capacitors with various gate electrodes and gate oxides K.-S. Chang-Liao and L.-C. Chen	942
Gas phase and surface reactions in subatmospheric chemical vapor deposition of tetraethylorthosilicate-ozone William N. Gill and Seshadri Ganguli.....	948
Testing of a rapid fault detection model for quality control: Borophosphosilicate glass thin films monitored by infrared absorption spectroscopy S. Zhang, J. E. Franke, T. M. Niemczyk, D. M. Haaland, J. N. Cox, and I. Banerjee.....	955
Highly preferred (111) texture aluminum-copper films formed with argon plasma treatment of the titanium underlayer and their electromigration endurance as interconnects Kazuyoshi Kamoshida and Yasuyuki Ito.....	961
Surface nitridation of silicon dioxide with a high density nitrogen plasma R. Kraft, T. P. Schneider, W. W. Dostalík, and S. Hattangady.....	967
Ar, N₂, and Cl₂ electron cyclotron resonance plasmas measured by time-of-flight analysis: Neutral kinetic energies and source gas cracking Rory S. Goodman, N. Materer, and Stephen R. Leone.....	971
Dry etch damage in GaAs metal-semiconductor field-effect transistors exposed to inductively coupled plasma and electron cyclotron resonance Ar plasmas F. Ren, J. W. Lee, C. R. Abernathy, S. J. Pearton, C. Constantine, C. Barratt, and R. J. Shul	983
Dry etching of germanium in magnetron enhanced SF₆ plasmas G. F. McLane, M. Dubey, M. C. Wood, and K. E. Lynch	990
Spatial variation of the etch rate for deep etching of silicon by reactive ion etching Bo Asp Møller Andersen, Ole Hansen, and Martin Kristensen.....	993
Efficiency evaluation of postetch metal stack anticorrosion treatments using chemical analyses by x-ray photoelectron spectroscopy and wide dispersive x-ray fluorescence P. Czuprynski, O. Joubert, M. Heitzmann, D. Louis, C. Vizios, and E. Lajoinie	1000

Rapid Communications

Isolation of a lattice-mismatched AlInAs/GaInAs layer on InP using ion implantation for high energy mobility transistor realization Hervé Fourré, Jean Claude Pesant, Olivier Schuler, and Alain Cappy.....	1008
Quantitative measurement of two-dimensional dopant profile by cross-sectional scanning capacitance microscopy J. S. McMurray, J. Kim, and C. C. Williams.....	1011

Papers from the 24th Conference on the Physics and Chemistry of Semiconductor Interfaces

Preface	1018
----------------------	------

Nanostructures and Self-Organization

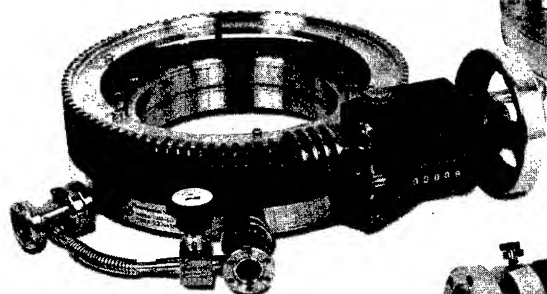
Atomic wires and their electronic properties Toshishige Yamada	1019
--	------

(Continued)

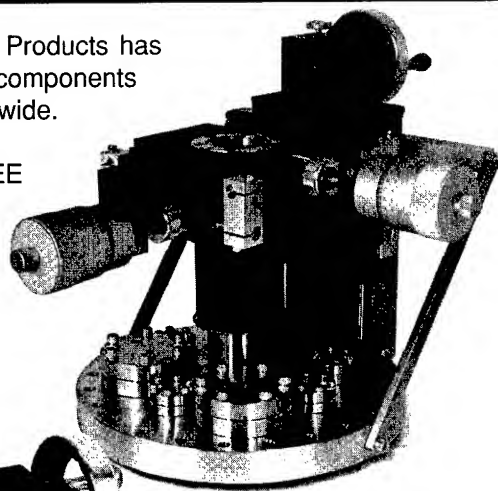
Proven Performance in UHV Manipulation

Since 1958, Thermionics Vacuum Products has provided the highest quality UHV components and systems to researchers worldwide.

Our exclusive 5-YEAR GUARANTEE attests our pride in workmanship and our commitment to your long-term satisfaction.



RNN™ Rotary Platforms



X,Y,Z Manipulators



Motion Feedthroughs

Our UHV specialties include:

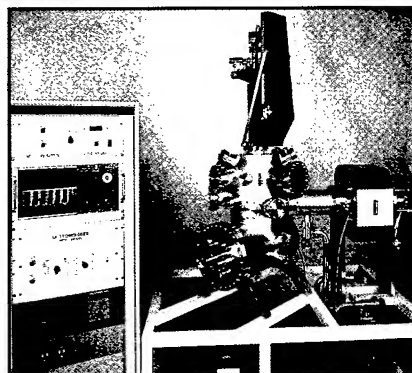
- ◆ Precision multi-axis motion
- ◆ Sample handling and transfer
- ◆ Rotary platforms
- ◆ Various heating and cooling options
- ◆ Motorized PLD target stages
- ◆ *In-situ* optical temperature measurement
- ◆ Non-clouding heated viewports
- ◆ Motorized, magnetically-driven rotary and linear feedthroughs



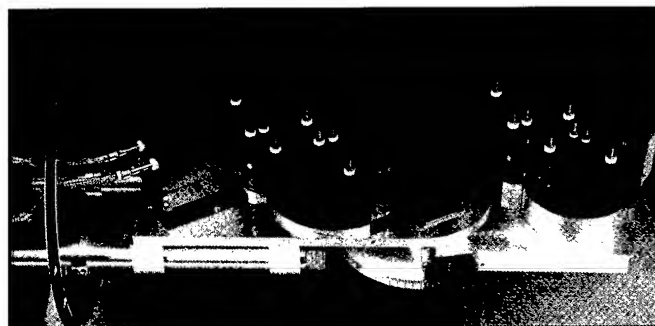
231-B Otto Street
Port Townsend, WA 98368

Phone: 360/385-7707 • FAX: 360/385-6617 • email: tnw@thermionics.com • <http://www.thermionics.com>

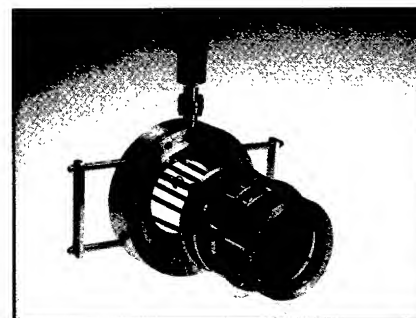
Surface Analysis Components and Systems Since 1985



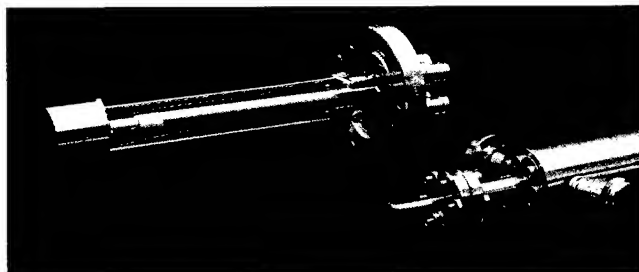
Custom UHV Systems



ELS3000 HREELS



LEED/Auger



NGI3000 ion guns and EG3000 electron guns



LK Technologies, Inc.
3910 Roll Avenue, Bloomington, IN 47403
Tel (812) 332-4449
Fax (812) 332-4493
<http://www.lktech.com>
e-mail: lktech@lktech.com

Morphological and compositional variations in strain-compensated InGaAsP/InGaP superlattices R. S. Goldman, R. M. Feenstra, C. Silfvenius, B. Stålnacke, and G. Landgren	1027
Spatial variations in luminescence and carrier relaxation in molecular beam epitaxial grown (InP)₂/(GaP)₂ quantum wires Y. Tang, D. H. Rich, A. M. Moy, and K. Y. Cheng	1034
Optical properties of semiconductor nanostructures T. L. Reinecke, P. A. Knipp, and S. N. Walck	1040
Electronic structure and optical behavior of self-assembled InAs quantum dots K. W. Berryman, S. A. Lyon, and Mordechai Segev	1045
Surface morphology and quantum dot self-assembly in growth of strained-layer semiconducting films K. E. Khor and S. Das Sarma	1051
Morphological stability of strained-layer semiconductors B. W. Wessels	1056
<i>Si Surfaces; Si/SiO₂ Interfaces</i>	
Second harmonic spectroscopy of Si(001) surfaces: Sensitivity to surface hydrogen and doping, and applications to kinetic measurements Z. Xu, X. F. Hu, D. Lim, J. G. Ekerdt, and M. C. Downer	1059
On the mechanism of the hydrogen-induced exfoliation of silicon M. K. Weldon, V. E. Marsico, Y. J. Chabal, A. Agarwal, D. J. Eaglesham, J. Sapjeta, W. L. Brown, D. C. Jacobson, Y. Caudano, S. B. Christman, and E. E. Chaban	1065
Minimization of suboxide transition regions at Si-SiO₂ interfaces by 900 °C rapid thermal annealing G. Lucovsky, A. Banerjee, B. Hinds, B. Claffin, K. Koh, and H. Yang	1074
Investigation of existing defects and defect generation in device-grade SiO₂ by ballistic electron emission spectroscopy H. J. Wen and R. Ludeke	1080
Macroscopic and microscopic studies of electrical properties of very thin silicon dioxide subject to electrical stress E. S. Daniel, J. T. Jones, O. J. Marsh, and T. C. McGill	1089
Plasma-assisted formation of low defect density SiC-SiO₂ interfaces A. Gözl, G. Lucovsky, K. Koh, D. Wolfe, H. Niimi, and H. Kurz	1097
<i>Si_{1-x}Ge_x and Related Materials</i>	
Raman spectra as a measure of interface alloying for IV/IV superlattices Jian Zi, Kaiming Zhang, and Xide Xie	1105
Measurement of band offsets in Si/Si_{1-x}Ge_x and Si/Si_{1-x-y}Ge_xC_y heterojunctions B. L. Stein, E. T. Yu, E. T. Croke, A. T. Hunter, T. Laursen, A. E. Bair, J. W. Mayer, and C. C. Ahn	1108
Interfaces of strained layer (Ge_nSi_m)_p superlattices studied by second-harmonic generation Xudong Xiao, Chun Zhang, A. B. Fedotov, Zhenghao Chen, and M. M. T. Loy	1112
<i>GaN and Other Wide-Bandgap Materials</i>	
Two-dimensional electron gas in AlGaIn/GaN heterostructures J. Z. Li, J. Y. Lin, H. X. Jiang, M. A. Khan, and Q. Chen	1117
Microstructure, vibrational and electronic properties of GaN grown by molecular beam epitaxy on Al₂O₃(0001) and 6H-SiC(0001) D. Freundt, D. Holz, H. Lüth, M. Romani, A. Rizzi, and D. Gerthsen	1121
Raman monitoring of molecular beam epitaxial growth of GaN on GaAs (100) and Si (111) D. Drews, A. Schneider, and D. R. T. Zahn	1128
Nitridation of the GaAs (001) surface using atomic nitrogen P. Hill, D. I. Westwood, L. Haworth, J. Lu, and J. E. Macdonald	1133
Optical transitions in InGaIn/AlGaIn single quantum wells K. C. Zeng, M. Smith, J. Y. Lin, H. X. Jiang, J. C. Robert, E. L. Piner, F. G. McIntosh, S. M. Bedair, and J. Zavada	1139
Theory of interfaces and surfaces in wide-gap nitrides Marco Buongiorno Nardelli, Krzysztof Rapcewicz, and J. Bernholc	1144
Growth and characterization of light emitting ZnS/GaN heterostructures E. C. Piquette, Z. Z. Bandić, J. O. McCaldin, and T. C. McGill	1148

(Continued)

Any Ordinary RGA Can Give You A Piece Of The Puzzle.



Any RGA (Residual Gas Analyzer) can give you partial pressure data. But the Dycor™ System 1000 gives you a total system for monitoring your vacuum process.

System 1000's input capabilities let you integrate parameters such as temperature, flow and pressure. So you can, for example, plot temperature and partial pressure in real time. And, outputs let you take corrective action if a process parameter exceeds your limits.

The most comprehensive system software in the industry gives you an intuitive graphical-user interface; incorporates an easy-to-use macro

With The Dycor™ System 1000 You Get The Total Picture.



language to automate any process at the touch of a button, controls all quadrupole operating parameters and stores data in the form most convenient for you.

Find out more about the Dycor System 1000, the RGA that gives you the total picture. Contact AMETEK, Process & Analytical Instruments Division, 150 Freeport Road, Pittsburgh, PA 15238. **Tel: 412-828-9040.**

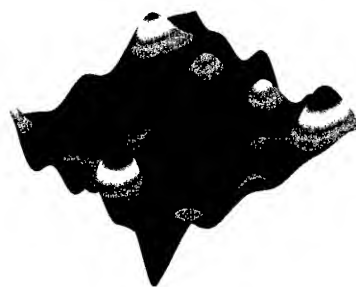
Fax: 412-826-0399.



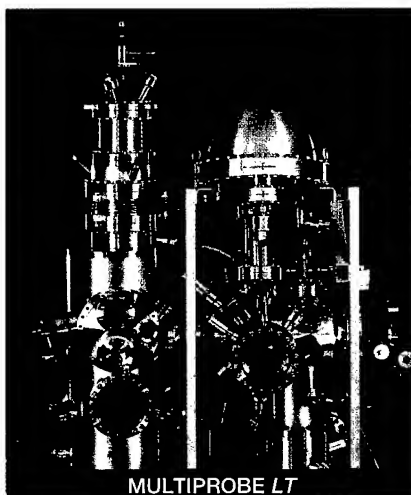
Low Temperature Surface Science

Combination of STM and Photoemission in the MULTIPROBE LT

STM

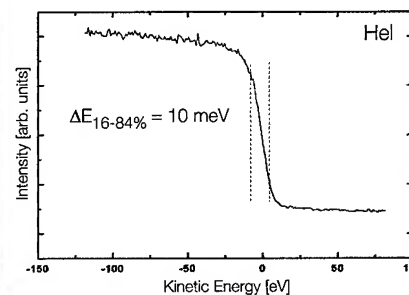


Single unit cell of a charge density wave on NbSe₂ at T = 5 K
(5 nm x 5 nm, U_{gap} = 0.02 V, I_t = 2.55 nA)



MULTIPROBE LT

UPS



High resolution electron spectroscopy at the Fermi edge of Ag at T = 10 K (sample mounted directly to the cryostat)



Headquarters:

OMICRON VAKUUMPHYSIK GmbH

Idsteiner Str. 78 • D-65232 Taunusstein • Germany
Tel. +49 (0) 6128 9 87 - 0 • Fax +49 (0) 6128 9 87 - 185

USA:

OMICRON ASSOCIATES

1226 Stoltz Road • Bethel Park, PA 15102
Tel. (412) 831-2262 • Fax (412) 831-9828

<http://www.omicron-instruments.com/>

Controlled formation of organic layers on semiconductor surfaces Jennifer S. Hovis, Seung Lee, Hongbing Liu, and Robert J. Hamers	1153
<i>III-V Heteroepitaxy and Passivation</i>	
Gallium desorption behavior at AlGaAs/GaAs heterointerfaces during high-temperature molecular beam epitaxy K. Mahalingam, D. L. Dorsey, K. R. Evans, and Rama Venkat	1159
Kink defects and Fermi level pinning on (2×4) reconstructed molecular beam epitaxially grown surfaces of GaAs and InP studied by ultrahigh-vacuum scanning tunneling microscopy and x-ray photoelectron spectroscopy Yasuhiko Ishikawa, Takashi Fukui, and Hideki Hasegawa	1163
Comparison of Si and GaAs/interfaces resulting from thermal and plasma oxidation P. R. Lefebvre and E. A. Irene	1173
Wet oxidation of AlAs films under ultrahigh vacuum conditions W. J. Mitchell, C.-H. Chung, S. I. Yi, E. L. Hu, and W. H. Weinberg	1182
Characterization of low-temperature grown AlSb and GaSb buffer layers K. G. Eyink, M. L. Seaford, T. W. Haas, D. H. Tomich, W. V. Lampert, S. D. Walck, J. S. Solomon, W. C. Mitchel, and L. F. Eastman	1187
Scanning tunneling microscopy and ballistic electron emission spectroscopy studies of molecular beam epitaxially grown Pt/CaF₂/Si(111) structures V. P. LaBella, L. J. Schowalter, and C. A. Ventrice, Jr.	1191
<i>Surface and Interface Spectroscopy</i>	
Evidence of near-surface localization of excited electronic states in crystalline Si L. Mantese, K. A. Bell, U. Rossow, and D. E. Aspnes.	1196
Microwave modulated photoluminescence as a contactless probe of interface states C. E. Inglefield, M. C. DeLong, P. C. Taylor, J. F. Geisz, and J. M. Olson.	1201
Surface and interface effects on ellipsometric spectra of crystalline Si K. A. Bell, L. Mantese, U. Rossow, and D. E. Aspnes.	1205
Reflectance-difference studies of interface-formation and initial-growth processes in ZnSe/GaAs(001) heteroepitaxy T. Yasuda, K. Kimura, S. Miwa, L. H. Kuo, A. Ohtake, C. G. Jin, K. Tanaka, and T. Yao.	1212
<i>Metal-Semiconductor Interfaces</i>	
Explanation of the linear correlation between barrier heights and ideality factors of real metal-semiconductor contacts by laterally nonuniform Schottky barriers R. F. Schmitsdorf, T. U. Kampen, and W. Mönch.	1221
Evolution mechanism of nearly pinning-free platinum/<i>n</i>-type indium phosphide interface with a high Schottky barrier height by <i>in situ</i> electrochemical process Hideki Hasegawa, Taketomo Sato, and Tamotsu Hashizume	1227
Comparison of electron affinity and Schottky barrier height of zirconium and copper-diamond interfaces P. K. Baumann and R. J. Nemanich	1236
<i>Heterojunctions and Heteroepitaxy</i>	
Nature and origins of stacking faults from a ZnSe/GaAs interface L. H. Kuo, K. Kimura, A. Ohtake, S. Miwa, T. Yasuda, and T. Yao	1241
Growth mode and defect generation in ZnSe heteroepitaxy on Te-terminated GaAs(001) surfaces A. Ohtake, L. H. Kuo, T. Yasuda, K. Kimura, S. Miwa, T. Yao, K. Nakajima, and K. Kimura	1254
<i>In situ</i> photoemission and reflectance anisotropy spectroscopy studies of CdS grown on InP(001) Ch. Schultz, A. M. Frisch, K. Hinrichs, J. Kinsky, T. Herrmann, U. Rossow, N. Esser, and W. Richter	1260
Effect of growth conditions on surface roughening of relaxed InGaAs on GaAs T. Pinnington, C. Lavoie, and T. Tiedje.	1265
In-rich 4×2 reconstruction in novel planar growth of InAs on GaAs(001) Q. K. Xue, Y. Hasegawa, T. Ogino, H. Kiyama, and T. Sakurai	1270
Subnanometer analysis of molecular beam epitaxy grown ternary arsenides M. L. Seaford, W. Wu, K. G. Eyink, D. H. Tomich, J. R. Tucker, and L. F. Eastman	1274

(Continued)

Over
7,000
Reasons to
Join **the**
American
Vacuum
Society

- ◆ *Receipt of 7,000+ pages of technical material*
- ◆ *Membership rates on conferences and symposia*
- ◆ *Discounts on short courses and educational materials*

All for under a penny a reason!
Our programs allow you to:

Interact with Technology Leaders at Symposia and Conferences

Each fall, scientists gather from around the world to attend the American Vacuum Society (AVS) National Symposium—a week-long forum for science and technology exchange in the areas of thin films, microelectronics, nanostructures, processing, surfaces, and vacuum. The symposium features papers from eight technical divisions and two technology groups, short courses, topical conferences on emerging technologies, and an equipment exhibition.

Local chapters host symposia on topics of interest to their local technical communities and equipment exhibitions during the year. Joining the AVS and a local chapter or division is an excellent way to interact with technology leaders and stay informed.

Increase Your Job Skills—More Than 70 Short Courses by Industry Experts

AVS short courses have provided hard-to-find job training for more than 20 years in applied vacuum technology, surface analysis and materials characterization, and materials, thin films, and coatings: processing and properties. The AVS offers over 70 short courses that have been developed by industry experts, and regional programs are held all year long to meet the training needs of local technical communities across the U.S.

Stay Informed—Receive Over 7,000 Pages of AVS Publications Each Year

AVS members receive the *Journal of Vacuum Science and Technology*—the first major physics journal available on CD-ROM. Refereed papers cover surface science, interfaces, nanometer-scale science, STM, plasmas, thin films, and vacuum technology. The *AVS Newsletter* announces meetings and events and includes interesting articles and conference reviews.

Surface Science Spectra is a unique laboratory reference tool published as a journal and on disk. It's an evolving database (including related sample and instrument information) of peer-reviewed XPS, Auger, and SIMS spectra of a wide range of materials.

Maintain Your Technical Edge with a Host of Educational Resources

The AVS has an ongoing program to develop and provide convenient, cost-effective educational and training resources. Formats include monographs (single-subject information); a series of book reprints entitled *AVS Classics*; various instructional materials; recommended practices information; and video tapes that provide industrial training, offer tutorial lectures, or present topical conferences.

Join Us Today—Under a Penny a Reason!

Membership dues are \$75 annually (\$20, students) and include receipt of *JVST*, *Physics Today*, and the *AVS Newsletter*; discounts on events and educational materials; and free membership in any of the AVS' chapters or divisions. (There is a separate subscription fee for *Surface Science Spectra*.)

The AVS is a volunteer-based, nonprofit technical society dedicated to advancing the science and technology of thin films, microelectronics, nanostructures, processing, surfaces, and vacuum.

For more information, contact the AVS, 120 Wall St., 32nd Floor, New York, NY 10005, 212-248-0200, fax 212-248-0245, e-mail avsnyc@vacuum.org, <http://www.vacuum.org>.



Local interface composition and extended defect density in ZnSe/GaAs(001) and ZnSe/In_{0.04}Ga_{0.96}As(001) heterojunctions

S. Heun, J. J. Paggel, L. Sorba, S. Rubini, A. Franciosi, J.-M. Bonard, and J.-D. Ganière 1279

Papers from the Fourth International Conference on Nanometer-Scale Science and Technology

Preface 1294

Surface Atomic Structure

Growth and characterization of Si and Ge clusters on ordered C₆₀ overlayers

D. Klyachko and D. M. Chen 1295

Interaction of C₆₀ with the (3×3) and ($\sqrt{3}\times\sqrt{3}$) surfaces of 6H-SiC(0001): Adsorption, decomposition, and SiC growth

L. Li, Y. Hasegawa, H. Shinohara, and T. Sakurai 1300

Novel C₆₀-DDME complex thin film with electrical bistable properties

M. Ouyang, Z. Q. Xue, K. Z. Wang, Q. D. Wu, and D. Qiang 1304

Si- and C-rich structure of the 6H-SiC(0001) surface

L. Li, Y. Hasegawa, and T. Sakurai 1307

Adsorption geometries of 1,9-decadiene on Si(111) 7×7 studied by scanning tunneling microscopy

D. Shachal and Y. Manassen 1310

Study on surface and interface structures of nanocrystalline silicon by scanning tunneling microscopy

Juning Gao, Haiqiang Yang, Ning Liu, Dongxia Shi, Yueshan Jiang, Zengquan Xue, Shijin Pang, and Yuliang He 1313

Diffusional attractions between voids on a Si(111)7×7 surface

E. Ter Ovanesyan, Y. Manassen, N. Ramesh Rao, and Z. Olami 1317

Liquid-phase adsorption process of β -picoline on stilbite (010) observed by atomic force microscopy

Masaharu Komiyama and Minming Gu 1325

Nanoscale Magnetic Properties

Magnetic nanostructures studied by scanning probe microscopy and spectroscopy

R. Wiesendanger, M. Bode, M. Kleiber, M. Löhndorf, R. Pascal, A. Wadas, and D. Weiss 1330

Magnetotransport through mesoscopic antidot arrays

Hongqi Xu 1335

Magnetic property of NiO ultrafine particles with a small Ni core

Sanshiro Sako, Kazunari Ohshima, Masahiro Sakai, and Shunji Bandow 1338

Magnetic force microscope images of magnetic domains in magnetic garnet

F. Tian, C. Wang, G. Y. Shang, N. X. Wang, and C. L. Bai 1343

Field induced antiferromagnetic ordering in NiS cluster confined in zeolite Y

Weiping Ding, Wei Zhong, Ning Zhang, Zaibing Guo, Gang Gu, and Youwei Du 1347

Study of the complex of the Schiff-base polymer with sulfate iron by magnetic force microscopy

F. Tian, C. Wang, C. L. Bai, W. G. Li, and M. X. Wan 1350

Imaging, polymerization, and reconstruction of polystyrene films with a scanning tunneling microscope

Z. Y. Hua and W. Xu 1353

Application of the scanning force microscope in structuring and in temperature-dependent analysis of Au nanostructures

H. Göbel, L. Jacobs, and P. von Blanckenhagen 1359

Surface Modification and Nanolithography

Scanning tunneling microscopy modification of Ag thin films on Si(100): Local rearrangement of the Si substrate by Ag/Si eutectic phase formation

U. Memmert, U. Hodel, and U. Hartmann 1364

Novel lithography and signal processing with water vapor ions

H. W. P. Koops, E. Dobisz, and J. Urban 1369

Fabrication of surface nanostructures by scanning tunneling microscope induced decomposition of SiH₄ and SiH₂Cl₂

H. Rauscher, F. Behrendt, and R. J. Behm 1373

(Continued)

Studies of field related effects in the fabrication process on graphite using a scanning tunneling microscope	
Chen Wang, Xiaodong Li, Guangyi Shang, Xiaohui Qiu, and Chunli Bai	1378
Nanostructuring by reactive accelerated cluster erosion	
A. Gruber, J. Gspann, and P. von Blanckenhagen.....	1382
Molecular patterns by manipulating DNA molecules	
Zhen-qian Ouyang, Jun Hu, Sheng-fu Chen, Jie-lin Sun, and Min-qian Li	1385
Observation of modification and recovery of local properties of polyethylene oxide	
H.-Y. Nie, M. Motomatsu, W. Mizutani, and H. Tokumoto.....	1388
Scanning probe nanofabrication of chemically active areas on substrate covered with organosilane monolayers	
Hiroyuki Sugimura and Nobuyuki Nakagiri.....	1394
Fabrication of hybrid superconductor-semiconductor nanostructures by integrated ultraviolet-atomic force microscope lithography	
Pasqualantonio Pingue, Marco Lazzarino, Fabio Beltram, Ciro Cecconi, Paolo Baschieri, Carlo Frediani, and Cesare Ascoli.....	1398
<i>Single Electron Tunneling</i>	
Metal-based single electron transistors	
W. Chen and H. Ahmed.....	1402
Electron beam dot lithography for nanometer-scale tunnel junctions using a double-layered inorganic resist	
Satoshi Haraichi, Toshimi Wada, Sucheta M. Gorwadkar, Kenichi Ishii, and Hiroshi Hiroshima.....	1406
Negative differential resistance on single electron transport in a junction array of ultrasmall islands	
Hiroshi Nakashima and Kiyohiko Uozumi.....	1411
<i>Self-Assembly and Self-Organization</i>	
Patterning of Langmuir-Blodgett film with ultrahigh vacuum-scanning tunneling microscope/atomic force microscope	
H. Hamanaka, T. Ono, and M. Esashi	1414
Molecular organization of azobenzene derivatives at the liquid/graphite interface observed with scanning tunneling microscopy	
P. C. M. Grim, P. Vanoppen, M. Rücker, S. De Feyter, S. Valiyaveetil, G. Moessner, K. Müllen, and F. C. De Schryver	1419
Formation of nanocolumn self-assembly by solvent polarity control	
Xiangdong Chai, Wensheng Yang, Yunwei Cao, Yueshun Jiang, Ran Lu, Tiejin Li, Huixin He, Yongqiang Wang, and Zhongfan Liu	1425
Writing and reading bit arrays for information storage using conductance change of a Langmuir-Blodgett film induced by scanning tunneling microscopy	
K. Takimoto, R. Kuroda, S. Shido, S. Yasuda, H. Matsuda, K. Eguchi, and T. Nakagiri	1429
Semiconducting, gas-sensing properties of Europium bisphthalocyanine Langmuir-Blodgett thin films	
Bingjie Liang, Chunwei Yuan, and Yu Wei.....	1432
<i>Nanowires and Nanoparticles</i>	
Formation of Si nanowire by atomic manipulation with a high temperature scanning tunneling microscope	
R. Hasunuma, T. Komeda, H. Mukaida, and H. Tokumoto.....	1437
Study on photoelectric properties of a TiO₂ nanoparticle	
Yongmei Chen, Yaan Cao, Yubai Bai, Wensheng Yang, Jihua Yang, Haiyan Jin, and Tiejin Li	1442
Formation of nanocrystals in <i>a</i>-Si thin films induced by pulsed laser ultraviolet irradiation	
Z. Chvoj, V. Cháb, and O. Borusík	1445
Atomic force microscopy investigations on the surface topographies of aluminum-based composite containing nanocluster diamond	
Q. Ouyang, B. Wang, and K. Okada.....	1449
Calculation of electronic energy levels in artificially confined cavities of a sphere and a circular cylinder	
Yong Fang Zhao, Xiao Gong Jing, Lin Song Li, Li Jun Wang, Zheng Hui, and Tie Jin Li	1452
Preparation and structural characterization of nanostructured CoAg granular films	
H. Sang, G. Ni, J. Lu, H. Chen, J. R. Zhang, and Y. W. Du.....	1456

(Continued)

Fabrication, characterization, and the photoelectric conversion of the nanostructured TiO₂ electrode Huihua Deng, Zuhong Lu, Haifang Mao, and Huijun Xu	1460
Surface vibrational mode of ZnS nanoparticles Jianfeng Xu, Haitao Mao, Yue Sun, and Youwei Du	1465
Photovoltaic study of nanocrystalline TiO₂ film modified with dye molecules Jinghuai Fang, Jingwen Wu, Xiaomei Lu, Jianhua Gu, and Zuhong Lu	1468
<i>Optical Responses</i>	
Visible luminescence from Si/SiO₂ superlattices S. V. Novikov, J. Sinkkonen, O. Kilpelä, and S. V. Gastev	1471
Numerical study of optical transmission in sandwich slabs: Implication to photon scanning tunneling microscopy Shu Wang, Mufei Xiao, and Jesús M. Siqueiros	1474
<i>Quantitative Topographic Analyses</i>	
Analysis of frictional-force image patterns of a graphite surface Naruo Sasaki, Masaru Tsukada, Satoru Fujisawa, Yasuhiro Sugawara, Seizo Morita, and Katsuyoshi Kobayashi	1479
Quantification of topographic structure by scanning probe microscopy James D. Kiely and Dawn A. Bonnell	1483
Error budget of step height and pitch measurement using a scanning tunneling microscope with a three-dimensional interferometer Toru Fujii, Masataka Yamaguchi, Masatoshi Suzuki, Hirofumi Yamada, and Kan Nakayama	1494
Principle of atomic grating and its application in nanotechnology H. W. Hao, H. Qiao, H. K. Cheng, W. W. Wei, and W. K. Yi	1498
Size and arrangement of elementary fibrils in crystalline cellulose studied with scanning tunneling microscopy Y. Z. Zhang, X. L. Chen, J. Liu, P. J. Gao, D. X. Shi, and S. J. Pang	1502
Nanomechanical surface characterization by atomic force acoustic microscopy U. Rabe, V. Scherer, S. Hirsekorn, and W. Arnold	1506
<i>Novel Instrumentation</i>	
Detection mechanism of an optical evanescent field using a noncontact mode atomic force microscope with a frequency modulation detection method Masayuki Abe, Takayuki Uchihashi, Masahiro Ohta, Hitoshi Ueyama, Yasuhiro Sugawara, and Seizo Morita	1512
Fabrication of probe tips for reflection scanning near-field optical microscopes: Chemical etching and heating-pulling methods Mufei Xiao, Jesús Nieto, Roberto Machorro, Jesús Siqueiros, and Hector Escamilla	1516
Near-field optical microscopy with uncoated tips: Calibration, chemical contrast on organic crystals, and photolithography G. Kaupp, A. Herrmann, and M. Haak	1521
Preparation of probe tips with well-defined spherical apexes for quantitative scanning force spectroscopy U. D. Schwarz, O. Zwörner, P. Köster, and R. Wiesendanger	1527
Fabrication of a Si scanning probe microscopy tip with an ultrahigh vacuum-scanning tunneling microscope/atomic force microscope Takahito Ono, Hiroaki Saitoh, and Masayoshi Esashi	1531
Conductive supertips for scanning probe applications C. Schössler, J. Urban, and H. W. P. Koops	1535
Simultaneous optical detection techniques, interferometry, and optical beam deflection for dynamic mode control of scanning force microscopy M. Houmady, E. Farnault, T. Yahiro, and H. Kawakatsu	1539
Development of ultrahigh vacuum-atomic force microscopy with frequency modulation detection and its application to electrostatic force measurement Takayuki Uchihashi, Masahiro Ohta, Yasuhiro Sugawara, Yoshio Yanase, Tatsuhiko Sigematsu, Mineharu Suzuki, and Seizo Morita	1543

(Continued)

Don't Lose Your Technical Edge ...

Attend the premier annual symposium on materials, surfaces, interfaces, and processes!

American Vacuum Society 44th National Symposium October 20-24, 1997, San Jose, California

Eight technical division programs will be held on issues related to films, microelectronics, nanostructures, processing, surfaces, and vacuum, including:

- ◆ Applied Surface Science.
- ◆ Electronic Materials and Processing.
- ◆ Nanometer-Scale Science and Technology.
- ◆ Plasma Science and Technology.
- ◆ Surface Science.
- ◆ Thin Films.
- ◆ Vacuum Metallurgy.
- ◆ Vacuum Technology.

In addition, the week-long annual symposium will feature:

Three Topical Conferences

- ◆ Flat Panel Displays (FPD).
- ◆ Micro-electromechanical Systems (MEMS).
- ◆ Quantitative Surface Analysis.

Three Technical Group Programs

- ◆ Biomaterial Interfaces.
- ◆ Magnetic Interfaces and Nanostructures.
- ◆ Manufacturing Science and Technology.

40+ Short Courses

- ◆ Applied Vacuum Technology.
- ◆ Surface Analysis and Materials Characterization.
- ◆ Materials, Thin Films, and Coatings: Processing and Properties.

150+ Exhibitors

- ◆ Vacuum and Deposition Equipment.
- ◆ Analysis Systems.
- ◆ Vacuum-Related Services.

For more information on the Symposium, visit the AVS Web site at <http://www.vacuum.org> or contact Angela Mulligan, (212) 248-0200, fax (212) 248-0245, e-mail avsnyc@vacuum.org.



The American Vacuum Society is a volunteer-based, nonprofit organization dedicated to advancing the science and technology of films, microelectronics, nanostructures, processing, surfaces, and vacuum and to providing a variety of educational opportunities. There are eight technology-based divisions, three technical groups, 21 local-area chapters, and about 6,000 members.

Development of a metal patterned cantilever for scanning capacitance microscopy and its application to the observation of semiconductor devices	
Takuma Yamamoto, Yoshihiko Suzuki, Masayuki Miyashita, Hiroyuki Sugimura, and Nobuyuki Nakagiri	1547
Novel high vacuum scanning force microscope using a piezoelectric cantilever and the phase detection method	
Jiaru Chu, Toshihiro Itoh, Chengkuo Lee, Tadatomo Suga, and Kazutoshi Watanabe	1551
New technique for nanocantilever fabrication based on local electrochemical etching: Applications to scanning force microscopy	
M. Hoummady, E. Farnault, H. Fujita, H. Kawakatsu, and T. Masuzawa	1556
Development of a piezoelectric self-excitation and self-detection mechanism in PZT microcantilevers for dynamic scanning force microscopy in liquid	
Chengkuo Lee, Toshihiro Itoh, Takahiro Ohashi, Ryutaro Maeda, and Tadatomo Suga	1559
Interface imaging by second-harmonic microscopy	
M. Flörsheimer, M. Bösch, Ch. Brillert, M. Wierschem, and H. Fuchs	1564
Scanning acoustic tunneling microscopy and spectroscopy: A probing tool for acoustic surface oscillations	
T. Hesjedal, E. Chilla, and H.-J. Fröhlich	1569
Investigation on a novel vacuum microelectronic pressure sensor with stepped field emission array	
S. H. Xia, J. Liu, D. F. Cui, J. H. Han, S. F. Chen, and L. Wang	1573
Robotic nanomanipulation with a scanning probe microscope in a networked computing environment	
C. Baur, B. C. Gazen, B. Koel, T. R. Ramachandran, A. A. G. Requicha, and L. Zini	1577
Using a new kind of organic complex system of electrical bistability for ultrahigh density data storage	
H. J. Gao, L. P. Ma, H. X. Zhang, H. Y. Chen, Z. Q. Xue, and S. J. Pang	1581
Atomic force microscope-based data storage using replicated media	
B. D. Terris, S. A. Rishton, H. J. Mamin, M. E. Best, J. A. Logan, and D. Rugar	1584
AUTHOR INDEX	1588
CUMULATIVE AUTHOR INDEX	1590

JOURNAL OF VACUUM SCIENCE AND TECHNOLOGY

INFORMATION FOR CONTRIBUTORS

1. Articles. This journal publishes two types of articles: original research reports comprising a complete, detailed, self-contained description of a research effort and Critical Review papers that provide an updating of an area of science or technology pertinent to the scope of this journal. All articles must contain an abstract (see item 5).

2. Brief Reports and Comments. These are short submissions that either report a brief item of research or technology that does not require extensive descriptive writing or that comment on an item previously published in the journal. Brief Reports and Comments do not require the format structure of regular journal articles and do not require an abstract.

3. Rapid Communications. These are short submissions which contain important new results and are intended for accelerated publication. Rapid Communications are given priority in handling and do not require an abstract.

4. Shop Notes. These are "how to do it" papers. They should be written and illustrated so that the reader may easily follow whatever instruction or advice is being given. An abstract is not required for a Shop Note.

5. Abstracts. An abstract must accompany each article. The abstract should be typed (double-spaced) on a separate page. It should be adequate as an index and as a summary. It should give all subjects, major and minor, concerning which new information is presented. It should give the conclusions of the article and all numerical results of general interest. An abstract is usually reprinted verbatim in abstract journals. Therefore, great care should be used in writing it.

6. Manuscripts. Manuscripts are accepted with the understanding that they have not been copyrighted, published, or accepted for publication elsewhere. Manuscripts must be in English, typewritten, double-spaced on one side of the page, on 8½×11 in. durable, opaque, white paper. *An original and two (2) copies must be submitted.* Wide margins should be provided to permit editorial instructions to the printer. *Authors should adopt a particular way of writing their names in the "by-line" and use the same name for all publications.* This practice makes indexes more useful and less confusing. Upon acceptance, a copy of the text portion of the manuscript on computer disk is requested. See item 15 below.

7. Style. Authors are expected to follow the conventional writing, notation, and illustration style prescribed in the American Institute of Physics *Style Manual* which is available from the AIP, c/o AIDC, 64 Depot Rd., Colchester, VT 05446; 1-800-809-2247, for \$10.00 prepaid. Authors should also study the form and style of printed material in this journal. The SI units should be used.

8. Illustrations: Illustrations published in the journal are either scanned by AIP using a digital scanner or received electronically from the author, and integrated with the text of the article, creating completely electronic pages. To receive optimal quality, we strongly encourage you to send electronic graphics files to AIP, rather than laser output. (Note: If you are submitting electronic graphics files, you are still required to send hardcopies of the figures to the Editorial Office. Adherence to electronic submission instructions is crucial. If your electronic files are received after AIP has already processed the hardcopy illustrations, the electronic files will not be used.) Please adhere to the following guidelines when preparing your illustrations for submission:

Sizing Illustrations (Electronic Graphics Files and Hardcopy)

- **Prepare illustrations in the final published size, not oversized or undersized.** Size your illustrations according to your journal's specifications. Submit each illustration at the final size in which it will appear in the journal. The standard is 8.5-cm maximum width (3-3/8" or 20.5 picas) for one column. This is especially important for screened or shaded illustrations; reduction of screened/shaded originals during the digitizing process introduces an unacceptable moiré pattern.
- **Ensure a minimum of 8-point type size (2.8 mm high; 1/8" high) and 1-point line width within illustrations.** Ensure that line weights will be 0.5 points or greater in the final published size. **Line weights below 0.5 points will reproduce poorly.** Avoid inconsistencies in lettering within individual figures, and from one figure to the next. Lettering and symbols cannot be handwritten. Avoid small open symbols that tend to fill in if any reduction is necessary.

Preparation of Hardcopy Illustrations for Scanning

- **Number figures in the order in which they appear in text.** Label illustrations with their number, the name of the first author, and the journal, on the front of the figure well outside the image area.
- **Place only one figure per page (including all parts).** Place all parts of the same figure on one sheet of white bond paper, spaced 1/4 in. apart, using a glue stick or wax on the back of the illustration and leaving a 2-in. bottom margin. Label all figure parts with (a), (b), etc. Make sure each figure is straight on the page. *Photocopies of artwork are not acceptable.*
- **Do not use correction fluid or tape on illustrations.** The scanner is extremely sensitive and reproduces all flaws (e.g., correction fluid, tape, smudges, dust). *Do not write on the back of the figure* because it will be picked up by the scanner.
- **Authors' laser-generated graphics are acceptable only if the lettering and lines are dark enough, and thick enough, to reproduce clearly,** especially if reduction is required. Maximum black-white contrast is necessary. Choose a laser printer with the highest dot-per-inch (dpi) available (i.e., the highest resolution possible). Remember that fine lines in laser-generated graphics tend to disappear upon reduction, even if the oversized original looks acceptable.
- **Submit continuous-tone photographs in final published size on white glossy or matte paper.** Avoid glossy paper stock that is off-white, ivory, or colored because contrast within the illustration will be lost in reproduction. Print the photograph with more contrast than is desired in the final printed journal page. Avoid dull, textured paper stock, which will cause illustrations to lose contrast and detail when reproduced.

Preparation of Electronic Graphics Files

- We recommend that all halftone art (screened art), shaded figures, and combinations (line art + halftone) be submitted electronically. Computer-generated illustrations output to desktop printers produce a screen. These figures are most problematic in the scanning process, because scanning screened output produces an unacceptable moiré pattern.
- **Acceptable formats:** Graphics must be submitted as PostScript, EPS (using either Arial or Times Roman fonts), or TIFF (lzw compressed). Do not send application files, e.g., Corel Draw, etc.
- **Settings:** Set the graphic for **600 dpi** resolution for line art, **264 dpi** for halftones (noncompressed), and **600 dpi** for combinations (line art + halftone). Save the files to grayscale (B/W), not color.
- Make sure there is only **ONE** figure per file. Each figure file should include all parts of the figure. For example, if Figure 1 contains three parts (a, b, c), then all of the parts should be combined in a single file for Figure 1.
- You are still required to send hardcopies of all figures to the Editorial Office, along with a hardcopy of the manuscript.
- Do not FTP the graphics files to the Editorial Office or AIP unless otherwise instructed to do so. Full instructions will be sent to you twice: Once on a hardcopy form after initial submission of your article, and once via e-mail after your article has been accepted for publication.

Detailed instructions for submitting electronic graphics to AIP and a glossary of terms may be found on the AIP Physics Information Netsite at www.aip.org/epub/submitgraph.html

9. Mathematics. Mathematical expressions should be typewritten as completely as possible, with unavailable symbols being carefully inserted in ink. Special care should be given to make equations and formulas clear to the typesetter, and all handwritten letters and symbols that cannot be typed should be identified in the margin and only the first time that such occur. Clearly distinguish capital and lowercase letters where there could be confusion. Use fractional exponents to avoid root signs. Extra symbols should be introduced to avoid complicated exponents or where it is necessary to repeat a complicated expression a number of times. Use the solidus (/) wherever possible for fractions. Do not repeat mathematical derivations that are easily found elsewhere in the literature; merely cite the references.

10. References. *References should be collected and typed (double-spaced) on a separate page.* A list of the abbreviations for the names of journals appears in the *Style Manual*. They should be numbered consecutively and arranged thus:

¹A. B. Smith, *Phys. Rev.* **41**, 852 (1932).

²H. Lamb, *Hydrodynamics*, 6th ed. (Cambridge University Press, Cambridge, England, 1940), pp. 573, 645.

11. Tables. All but the simplest tabular material should be organized into separate tables. Tables should be numbered with Roman numerals and typed on sheets at the end of the running text. Each table must have a caption typed at the top of the table, which makes the data in the table intelligible without reference to the text. Avoid complicated column headings. If necessary, use symbols which are explained in the caption. Type a double horizontal line below the caption, a single line below the headings, and another double line at the end of the table. Long tables should, if possible, be submitted in a form ready for direct photoreproduction. Detailed instructions for preparing camera-ready tables are available from the Editorial Office.

12. Publication Charges comprise the major source of income for this journal. Consequently, author's institutions or companies are asked *and urged* to pay a publication charge—which entitles them to 100 reprints without covers—of \$95 per printed page.

13. Correspondence and Alterations. *Send manuscripts to the Editorial Office, Journal of Vacuum Science and Technology, Microelectronics Center of North Carolina, 3021 Cornwallis Rd., Caller Box 13994, Research Triangle Park, North Carolina 27709-3994; telephone: (919) 248-1861; FAX: 919-248-1857; E-mail: jvst@mcnc.org.* Manuscripts to be published in special issues (papers from conferences) should be sent to the Guest Editor appointed for the particular Conference. Please do not send the Editor any correspondence about proofs, reprints, or publication charges.

If your paper is accepted for publication, you will be notified by the Editor. Any alterations can be made at this time. You will also be notified when

your paper is sent to the American Institute of Physics. *All subsequent correspondence about your paper should be addressed to the Editorial Supervisor, JVSTA or JVSTB, American Institute of Physics, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999. Reference must be made to the author, journal, and scheduled date of issue.*

14. Proofs of Articles will be sent to you and should *be returned promptly to the Editorial Supervisor, JVSTA or JVSTB, AIP.* A few alterations in proof are unavoidable, but the cost of making extensive alterations or of correcting mistakes caused by careless preparation of the manuscript will be charged to the author. *Unless corrected proofs are returned punctually, your paper will be published in uncorrected form and will be so labeled.*

15. Computer Disks. A copy of the final, accepted manuscript on a computer disk should be submitted to the Editorial Office, JVST, MCNC, 3021 Cornwallis Rd., Caller Box 13994, RTP, NC 27709. Disks must be labeled with the manuscript file name, manuscript title and authors, and word processing format used. The requested format is WordPerfect or Microsoft Word, but disk versions may also be submitted in ascii, L^AT_EX, or REVTeX formats. It is imperative that the guidelines as documented in either the MS Word/WordPerfect toolbox or the REVTeX toolbox be followed precisely. If electronic retrieval is not possible, you may receive the toolbox or toolkit on disk by contacting toolkits@aip.org. The REVTeX toolbox is available via anonymous FTP on the Internet from <ftp.aip.org> in the directory `/pub/revtex`. The MS Word/WordPerfect toolbox is also available via anonymous FTP from <ftp.aip.org>. Move to the directory `/ems`, then follow the instructions given on the screen. The REVTeX toolbox and the Word/WP toolkit, as well as general information regarding the compuscript program, are accessible via the AIP Physics Information Netsite (www.aip.org/epub/compuscripts.info.html). The Word/WP author toolkit FAQ ("Frequently Asked Questions") may be found at www.aip.org/aip/toolkit.html.

16. Electronic Mail. Authors may include their internet e-mail addresses along with all other footnotes in the following format: Electronic mail: smith@foo.bar.com

Microstructure control in semiconductor metallization

J. M. E. Harper^{a)} and K. P. Rodbell

IBM T. J. Watson Research Center, Yorktown Heights, New York 10598

(Received 24 April 1997; accepted 25 April 1997)

The microstructure of semiconductor metallization is becoming increasingly important as linewidths decrease below $0.5\ \mu\text{m}$. At these dimensions, reliability and performance are greatly influenced by specific microstructural features rather than only by the average material properties. In this article, we address the prospects for controlling the microstructure of thin film interconnection metals as linewidths are predicted to decrease below $0.1\ \mu\text{m}$ by the year 2010. First, we evaluate the sources of energy available to drive microstructure changes in thin films, both during and after deposition. The internal energy sources considered are grain boundaries, interfaces, surfaces, strain, solidification, crystallization, solute precipitation, and phase transformations, with energy densities ranging from less than $1\ \text{meV/atom}$ to greater than $100\ \text{meV/atom}$. The external energy sources considered are particle bombardment during deposition, mechanical deformation, and radiation damage, which may deliver energies greater than $100\ \text{eV/atom}$. Second, we review examples of microstructure changes in terms of these energy sources. These examples include the dependence of Al-Cu and Ti fiber texture on the roughness of SiO_2 , orientation change and abnormal Cu grain growth coupled to the precipitation of Co in Cu-Co alloys, and in-plane orientation selection during phase transformation of TiSi_2 in very narrow lines. A substantial degree of microstructure control is also achieved in films deposited with off-normal incidence energetic particle bombardment, which has been used to produce both in-plane and out-of-plane crystallographic orientations in metals (Mo, Nb), nitrides (AlN), and oxides (ZrO_2). Drawing on these examples, we discuss the prospects for microstructure control in future semiconductor metallization with respect to the list of energy sources, the decreasing dimensions, and the changing fabrication processes. One mechanism in particular, discontinuous precipitation of supersaturated solute atoms, is highlighted as having a substantial amount of stored energy available to drive microstructure evolution, and may provide a means to more fully control the microstructure of semiconductor metallization. © 1997 American Vacuum Society. [S0734-211X(97)09904-6]

I. INTRODUCTION

The Semiconductor Industry Association (SIA) predicts¹ that the minimum width of interconnection wiring in semiconductor metallization will decrease below $0.25\ \mu\text{m}$ by the year 2000, and below $0.1\ \mu\text{m}$ by the year 2010, as shown in Fig. 1. Also, the number of metal levels and the chip size will continue to increase, while the metal aspect ratio (height/width) is predicted to reach 4:1 by the year 2010 (Fig. 1). Even if these trends are not followed exactly, the need to gain control over the microstructure of conductor materials is still becoming more important. In current silicon interconnection structures, shown schematically in Fig. 2, the metal wiring levels (M1, M2, etc.) are fabricated from an Al-Cu alloy (typically Al-0.5 wt %Cu) embedded in SiO_2 insulation using a process sequence of blanket metal deposition followed by lithographic patterning, metal reactive ion etching (RIE), and SiO_2 deposition.² Multilevel interconnection structures include additional materials, shown in Fig. 2, such as tungsten interlevel studs and metal layers above and below the Al-Cu wiring. These layers may interact with the

Al-Cu conductor, forming, for example, Al_3Ti from the reaction of Ti "over/under" layers with the Al. The microstructure of the Al-Cu alloy films is largely determined by the process of deposition onto a planar surface. Typically, sputter deposited Al-Cu alloy films have a monomodal grain size distribution, with a strong crystallographic orientation in which (111) planes are parallel to the substrate surface, referred to as (111) fiber texture.³ Upon annealing to $400\ ^\circ\text{C}$, a common procedure in device fabrication, the grain sizes increase, while the grain size distribution remains monomodal, characteristic of "normal" grain growth.⁴ The (111) texture also becomes more well defined and a strong (111) texture has been shown to be important in determining the reliability of Al-Cu metallization patterned by RIE. For example, the electromigration lifetime of Al-Cu metallization increases as the strength of the (111) texture increases.⁵⁻⁷ It has also been shown that electromigration failure sites (hillocks and voids) are associated with grains that differ substantially from (111) orientation, and that the misorientation dependence of grain boundary diffusivity is an important parameter.⁸

As the semiconductor industry is moving towards using

^{a)}Electronic mail: harperj@watson.ibm.com

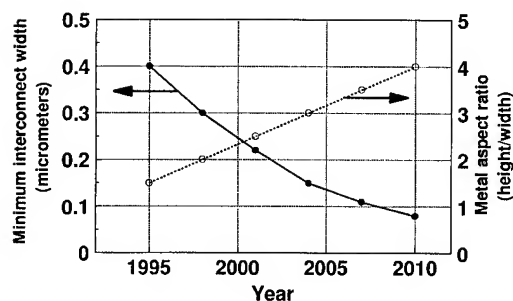


FIG. 1. Chart of the trends in minimum interconnect width and aspect ratio in semiconductor metallization from 1995 to 2010 (Ref. 1).

copper interconnection wiring for decreased resistivity,⁹ the crystallographic texture of Cu thin films has been examined and was found to have a more complex behavior than that of Al-Cu.^{4,10} With Cu thin films, the texture is not always strongly (111), and the presence of (200) or (220) oriented grains and twins is frequently observed.^{4,10} Also, the grain size distribution is often bimodal, with evidence of abnormal grain growth occurring during annealing, during which some grains grow substantially larger than the average grain size.^{4,11,12} Copper alloys have shown some unusual effects including dramatic grain growth and texture changes upon annealing at temperatures well below 400 °C.¹³ As a further complication, the process for fabricating copper interconnections differs greatly from that for Al-Cu since there is no acceptable RIE process available for Cu. The process being developed for Cu interconnections, referred to as the damascene process, includes masking and etching the insulator, followed by Cu deposition into the patterned insulator and removal of the excess Cu by chemical-mechanical planarization (CMP). A typical structure (Fig. 3) includes Cu wiring levels (M1, M2, etc.) with underlying liner materials containing metals such as Ti, Ta, W, or their nitrides to provide adhesion and diffusion barrier properties.¹⁴ These properties are important as copper is integrated with new interlevel insulators such as polyimides (shown as PI in Fig. 3). A "single damascene" structure is illustrated in Fig. 3, in

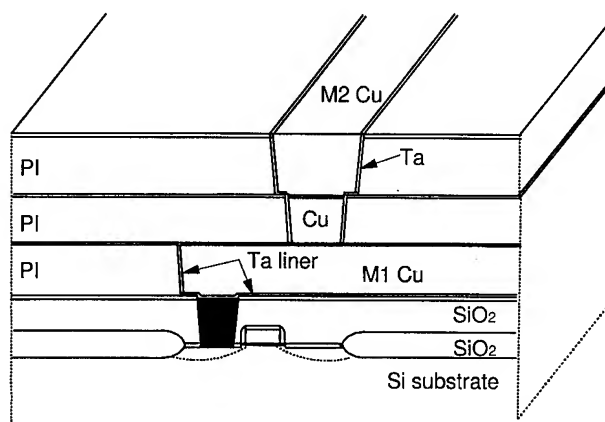


FIG. 3. Schematic diagram of a multilevel interconnection structure on a CMOS device using Cu metallization in polyimide (PI) insulation.

which a Cu interlevel stud is formed as a separate process step. Also under development is the "dual damascene" structure, in which both the stud and line levels are formed in one combined process step. In these damascene interconnection systems, the microstructure of Cu conductors is established within a patterned structure instead of on a planar surface. Damascene processing of Al-Cu is also being rapidly developed, and will produce microstructures different from current RIE processes that will need to be identified and controlled. The combined effects of these fabrication differences and their resulting microstructures on the performance and reliability of Al-Cu and Cu interconnections are not yet fully established, but it is very likely that additional control of the microstructure will be needed in the future to obtain the uniform material properties desired.

One method of influencing the microstructures of thin films is to deposit them onto underlayer materials with various textures or surface energies. At one extreme is epitaxial growth, in which a single crystal film is deposited with a crystallographic orientation completely determined by the single crystal underlying substrate. At the other extreme is a nonwetting film/substrate combination, in which the film has minimal interaction with the substrate, and its microstructure is determined by its own surface energies, which may cause a specific orientation to be selected during coalescence or growth.¹⁵ More typical, there is some interaction between the film and substrate leading to a preferred orientation of the film texture by minimization of this interface energy, as will be discussed below. Here, we will primarily consider examples in which a polycrystalline metal film is deposited onto an amorphous or polycrystalline substrate or underlayer material, such as Al-Cu on SiO₂ or Ti, and Cu and its alloys on SiO₂ or Ta. During heat treatment to 400 °C, the temperature exceeds one-third of the absolute melting temperature of the Al-Cu or Cu films, providing sufficient thermal energy for grain growth and offering an opportunity to control the resulting microstructure. We discuss the sources of energy available to control thin film microstructure evolution during annealing, including grain growth, interface energy minimization, strain energy minimization, solidification, crystallization, compound formation, solute precipitation, and phase

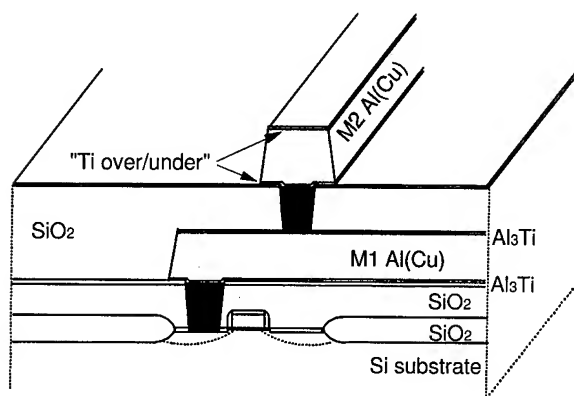


FIG. 2. Schematic diagram of a multilevel interconnection structure on a CMOS device using Al-Cu alloy metallization in SiO₂ insulation.

TABLE I. Sources of energy to control thin film microstructure.

Mechanism	Example	MJ/m ³	meV/atom
Grain growth	0.1–1.0 μm grain growth in Cu	15	1.1
Thermal energy	20–400 °C	300–700	25–60
Interface energy minimization	1.0- μm -thick Cu	2	0.15
Strain energy content	300 MPa biaxial stress in Cu	300	25
Strain energy minimization	(100) vs (111) Cu	1	0.1
Solidification	Cu	1500	110
Crystallization	Metals	1000	80
Compound formation	Al+Ti reaction to form Al ₃ Ti	800	80
Solute precipitation	Cu (4.9 at. % Ag)	90	7
Phase transformation	C49–C54 TiSi ₂	600	50
Bombardment during deposition	Ions or neutrals	$\sim 10^4$ – 10^6	$\sim 10^3$ – 10^5
Mechanical deformation	Cold rolling of Cu to >50% deformation	120	10
Radiation damage	Amorphization of Si or NiAl	$\sim 10^4$	$\sim 10^3$

transformation. External sources of energy such as mechanical deformation and radiation damage from ion implantation are also considered.

Depending on the properties to be optimized, a desirable microstructure might be a distribution of large grains with a strong fiber texture and a limited range of grain-to-grain misorientations, with possibly a preferred in-plane orientation, or it might be a distribution of small grains with a small variation in size, but a larger range of orientations. Here, we focus on the mechanisms that allow microstructure changes to occur, without specifying the final microstructure. Using the example of Cu–Co alloys, we highlight discontinuous precipitation of supersaturated solute atoms, in which solute precipitation is coupled to grain growth. We show that this mechanism contains a substantial amount of stored energy available to drive microstructure change, which may potentially be harnessed to control the microstructure of semiconductor metallization. We also discuss these mechanisms with respect to the trend towards high aspect ratio features patterned by damascene processing, for which the interfaces at the sidewalls of the conductors play an increasingly important role in determining the microstructure.

Although it is possible to change the microstructure of Al or Cu or their alloys by annealing at 400 °C, the underlayer materials typically have much higher melting temperatures, and their microstructures will not change at 400 °C. If we are to fully control the interconnection microstructure, then we must also consider how to influence the microstructure of refractory metal thin films during deposition. One method of controlling the microstructure of high melting temperature thin films is by using energetic particle bombardment during film deposition. There are many sources of energetic particle bombardment, including energetic ions of the sputtering gas,

energetic atoms of the sputtering gas reflected from the sputtering target, separate ion beam sources, and the energetic depositing atoms themselves. We will give examples of Nb, Mo, AlN, and ZrO₂ film deposition that illustrate several of these sources of energetic particles, and that demonstrate that both out-of-plane and in-plane orientations may be generated by the appropriate level of bombardment during deposition. Taken together, the methods for modifying crystallographic orientations both during and after deposition provide a range of opportunities for microstructure control that justify further exploration. We also raise some of the outstanding issues that will determine whether it is possible to gain full control of the microstructure of future semiconductor metallization.

II. SOURCES OF ENERGY TO CONTROL THIN FILM MICROSTRUCTURE

The processes of grain growth and microstructure evolution are often discussed in terms of driving forces responsible for the growth of certain grains or orientations.^{16,17} This approach is necessary to determine the kinetics of microstructure evolution, but a useful comparison of mechanisms may also be made by considering the sources of energy available to drive microstructure changes.^{18,19} We focus on interconnection metals in the thickness range of 0.1–1.0 μm , and estimate the energy density for 11 mechanisms that can influence thin film microstructures. These are tabulated in Table I, and presented graphically in Fig. 4. Also included in the list for comparison is the thermal energy per atom for the temperature range 20–400 °C. Even though thermal energy does not itself introduce anisotropic microstructures, it does activate many of the selection mechanisms listed below.

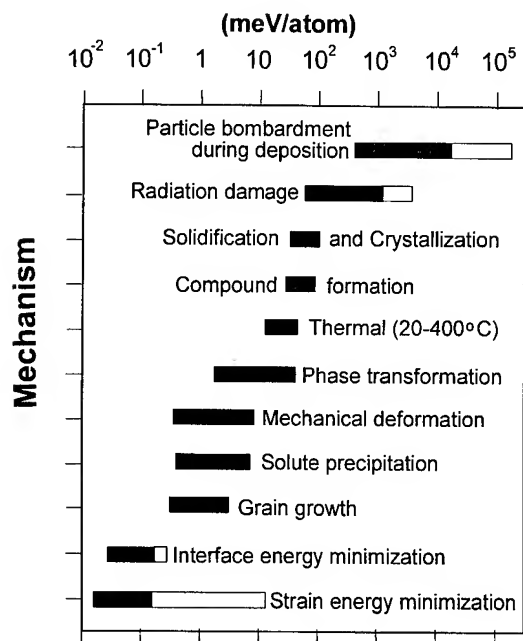


FIG. 4. Sources of energy (meV/atom) available to control thin film microstructure. The examples are also listed in Table I.

A. Grain growth

Depending on the ratio of the absolute substrate temperature to the absolute melting temperature of the metal (the homologous temperature T_s/T_m), the as-deposited grain structure is either "frozen" into a finely grained microstructure with grains of typically 10 nm in size, or is able to evolve during deposition to columnar or recrystallized microstructures of larger grain size up to a maximum of about twice the film thickness.²⁰ With a base pressure of about 10^{-9} – 10^{-8} Torr in the sputtering equipment typically used in semiconductor manufacturing, the boundary between these two regimes is at a homologous temperature of about 0.3, and decreases with decreasing impurity levels in the deposition environment.²¹ In such a sputtering environment, the grain size of Al and Cu films deposited at room temperature is smaller than the grain size obtained by annealing at 400 °C. During heating, curvature-driven grain boundary motion occurs, leading to the elimination of small grains and an overall increase in the grain size distribution. When grain growth occurs, shown schematically in Fig. 5, the energy released is given by the product of the grain boundary energy γ_B and the area of the eliminated grain boundaries.¹⁸ An estimate of the magnitude of this energy release is given by assuming that the starting grains are cubes of edge D , and the final grain size is much larger, in which case the energy is dominated by the starting grain size, and may be expressed as $3\gamma_B/D$. For copper, the high angle grain boundary energy γ_B is approximately 0.5 J/m^2 .^{16,18} If we assume that a Cu thin film is deposited with a starting grain size of $0.1 \mu\text{m}$, then subsequent grain growth up to a grain size of $1.0 \mu\text{m}$ releases an energy density of about 15 MJ/m^3 or 1.1 meV/atom . For comparison, the magnitude of thermal energy at

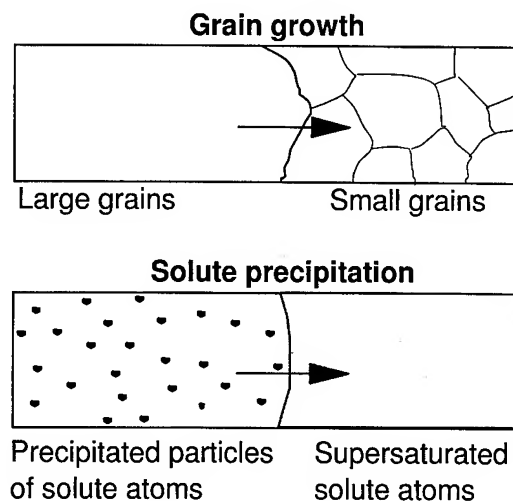


FIG. 5. Schematic diagram of boundary motion during grain growth and solute precipitation.

400 °C is about 60 meV/atom (Fig. 4). Therefore, on heating from 20 to 400 °C, the energy released during grain growth (from 0.1 to $1.0 \mu\text{m}$) is substantially less than the thermal energy.

B. Interface energy minimization

The strong (111) fiber texture usually obtained in sputtered aluminum alloy thin films is caused by the minimization of interface energies at the free surface and at the substrate interface.¹⁷ This energy term is especially significant during the early stages of film growth when island coalescence occurs, because it determines the initial texture of the growing film. If the growth environment is clean enough, subsequent film growth continues to replicate this texture.²¹ Only if the surface is subject to contamination within the monolayer deposition time and atomic mobility is low (homologous temperature below 0.3) does continuous nucleation of new grains occur, and a random orientation is possible.²⁰ Since the interface energy differences which distinguish grains are only present at the top and bottom surfaces, this energy term is not exactly an energy per unit volume. However, an estimate of its importance may be made by assuming a film thickness of $1.0 \mu\text{m}$ and calculating the magnitude of the interface energies per unit film volume for this thickness. The surface energy of Cu is about 1.7 J/m^2 ,¹⁶ and interface energies are comparable to grain boundary energies, about 0.5 J/m^2 . The net interface energy including the top and bottom interfaces is therefore about 2 J/m^2 , corresponding to 2 MJ/m^3 for a $1.0\text{-}\mu\text{m}$ -thick film, or 0.15 meV/atom . In Fig. 4, the overall magnitude of interface energy is indicated by the full length of the shaded and unshaded regions, and the shaded region indicates that some fraction of this energy, perhaps up to several tens of percent, is available to distinguish between grains of different orientations and to provide the selection mechanism for texture evolution. Clearly, if the film thickness is only $0.1 \mu\text{m}$, the energy density available from interface energy minimization

is 10 times higher per unit volume, and can therefore dominate texture evolution. This is one example of a source of energy that becomes increasingly important in very small dimensions, where the ratio of the number of atoms at interfaces to the number of atoms in the interior film volume is significant.

C. Strain energy minimization

Interconnection metal films may have substantial intrinsic stress as deposited, and will also develop a thermal mismatch stress upon heating; this is caused by the difference in thermal expansion coefficients between the metal film and the silicon substrate. For materials with high melting temperatures, such as refractory metals, the intrinsic stress is strongly affected by deposition conditions and can exceed the thermal mismatch stress. For lower melting temperature materials, such as Al and Cu, the intrinsic stress is typically smaller than the thermal mismatch stress. The thermal strain ϵ is given by the product of the temperature excursion ΔT and the difference in thermal expansion coefficients $\Delta\alpha$, and the induced film stress is the product of the biaxial modulus and the strain. For a typical stress of 300 MPa for Cu on Si ($\Delta\alpha = 13.6 \text{ ppm}/^\circ\text{C}$) and a temperature excursion of 300°C , the stored energy density is 300 MJ/m^3 , or about 25 meV/atom . Biaxial stress may be relieved by dislocation glide or other plastic deformation mechanisms, and the dimensional changes may cause delamination or metal crack formation. We list this energy content in Table I as "strain energy content" for comparison with the other mechanisms, and include it in Fig. 4 as the total length of the shaded and unshaded region labeled "strain energy minimization." The anisotropic nature of strain can strongly influence film microstructure. Due to the crystallographic anisotropy of elastic constants, grains of different orientations develop different levels of strain energy density as the film is placed under thermal mismatch stress during annealing. The result is that orientations with lower strain energy density are favored when grain growth occurs in the presence of stress, which may be locally nonuniform.^{17,22,23} The strain energy density difference between (100) and (111) oriented grains in Cu is given by $146.4\epsilon^2 \text{ GJ/m}^3$.²⁴ Applying this value to Cu on Si and a temperature excursion of 300°C , the strain energy difference is about 1 MJ/m^3 , corresponding to about 0.1 meV/atom . This term is shown in Fig. 4 as the shaded portion of strain energy minimization, and its magnitude is independent of whether the grains are under compressive or tensile stress.²⁴ The grain orientation in Cu favored by strain energy minimization differs from that favored by interface energy minimization. For conditions that favor strain energy minimization, i.e., large strain, thickness, and temperature excursions, the selected orientation in anisotropic fcc metals, such as Cu, is (100).¹⁷ For conditions of small strain, thinner films, and small temperature excursions, the favored orientation is selected by interface energy minimization, and is (111) in fcc metals. The transition between these two conditions has been modeled and demonstrated experimentally with initially randomly oriented Ag thin films on Ni on

MgO.²⁵ If plastic deformation occurs upon heating, the grain orientation that has the lowest yield stress (onset of dislocation glide) is selected. This criterion predicts a (110) texture, which has been observed in Al films.²⁵ Given this combination of mechanisms, the favored orientation is not always predictable, but the energy available from strain energy minimization can clearly be comparable to that available from grain growth.²²

D. Solidification

Current semiconductor processing does not require melting and solidification of the interconnection metal, but techniques are being developed to fill high aspect ratio features in which the metal is reflowed under conditions of high surface mobility (elevated temperature and ultrahigh vacuum environment).^{26,27} For the purpose of comparison with other sources of energy, we note that the latent heat of fusion is substantially higher than the energy content of the mechanisms listed above. For copper, this value is 1.5 GJ/m^3 , equivalent to about 110 meV/atom .¹⁸ Clearly, the microstructure is completely changed by melting and solidification, and new features have been observed when solidification takes place in recessed features with aspect ratios greater than unity. An example is the observation of a high concentration of resolidified Al-Cu grains which have (111) planes parallel to the sidewalls of a deep trench structure, instead of parallel to the substrate plane.²⁸ In these trenches, the sidewall area exceeds the bottom area, and a simple fiber texture is no longer obtained.

E. Crystallization

Closely related to solidification is crystallization from the amorphous state. Here, also, is a substantial source of energy for microstructure change, on the order of 1 GJ/m^3 , or 80 meV/atom (Ref. 29) (Fig. 4). Crystallization releases sufficient energy density in some cases to cause a self-propagating explosive process in which the crystallization front moves rapidly through the material from a point of initiation.³⁰ A large stress change in the tensile direction also accompanies this transformation because most materials experience a volume decrease upon crystallization.³¹ As with solidification, crystallization is not usually a part of present interconnection technology, and metals such as Al and Cu are not easily amorphized without a substantial concentration of additives. This source of energy might, however, be considered in the future for driving microstructure changes in metals.

F. Compound formation

An example of intermetallic compound formation in present semiconductor processing is the reaction of a Ti underlayer with deposited Al-Cu metallization upon heating to $300\text{--}400^\circ\text{C}$ to form Al_3Ti (Fig. 2). This compound is beneficial in extending the electromigration lifetime, but it removes some of the low-resistivity Al-Cu from the cross-sectional area of the conductor, and therefore the thickness of

the Al_3Ti must be minimized. Typical advanced complementary metal-oxide-semiconductor (CMOS) technology uses a deposited Ti thickness of 10–25 nm under an Al–Cu thickness of 300–500 nm, leading to a reacted Al_3Ti thickness of 30–80 nm.² The energy released in the Al–Ti reaction is given by the heat of reaction, approximately 30 kJ/g atom,³² which corresponds to 800 MJ/m³ of Al_3Ti volume or 80 meV/atom (Fig. 4). The available energy is somewhat reduced from this value by the initial interdiffusion of Al into Ti prior to compound formation.³² When this compound forms under an Al–Cu layer that is 5–10 times thicker than the Al_3Ti layer, the energy released averaged across the film thickness is correspondingly reduced, but is still comparable to the thermal energy (Fig. 4). This source of energy is not presently considered significant in affecting microstructure in Ti/Al–Cu/Ti layered metallization because most of the Al–Cu grain growth occurs at temperatures below the onset of the Al–Ti reaction.³² In the case of Cu metallization, some underlayers are being used (e.g., Ta) that do not react with Cu to form intermetallic compounds, making this source of energy unavailable unless other materials are included. It should be noted that the heat released in compound formation is capable of self-propagating explosive reactions if the number of interfaces is high within the overall film thickness, as was demonstrated in multilayer Ti/Si samples for which the propagation speed of the reaction front was measured directly.³³

G. Solute precipitation

Since sputter deposition involves a rapid quench from the vapor phase, metal films are easily deposited with solute atom concentrations well above the equilibrium concentration at the deposition temperature. During the first heating above the deposition temperature, the solute atoms gain enough mobility to precipitate into second phase elemental particles or to react with the matrix metal and precipitate as particles of a second phase compound. Solute precipitation is facilitated at grain boundaries, where the atomic structure is more open than in the grain interior, so the motion of a grain boundary through a supersaturated solution accelerates the release of the heat of solution, as shown schematically in Fig. 5. The result is a coupling of grain growth and solute precipitation referred to as discontinuous precipitation.¹⁸ The removal of atoms from solution releases an energy corresponding to the free energy difference between the metastable starting state with dissolved solute atoms and the final state with precipitated solute atoms. The magnitude of this energy change is primarily determined by the entropy difference between the configurations of dissolved versus precipitated solute atoms. This entropy difference depends primarily on the solute atom concentration c , and is approximately independent of the solute atom type if the solubility is near zero after precipitation occurs.³⁴ The generally exponential temperature dependence of the typical solubility curve $c(T)$ allows the entropy change to be expressed as a driving force, or energy density, of $R(T_1 - T_0)(c)\ln(c)$, where R is the gas constant, T_1 is the temperature at which the solute can be

brought into solution, and T_0 is the temperature of precipitation.¹⁸ An estimate for the energy released at 300 °C during discontinuous precipitation of 4.9 at. % Ag in Cu, which can be dissolved at 780 °C, is 90 MJ/m³, equivalent to 7 meV/atom (Ref. 18) (Fig. 4). Note that the sample does not actually need to be heated to 780 °C as long as the solute is initially deposited in supersaturated solution. This energy density is substantially higher than those contained in the mechanisms of interface energy minimization and strain energy minimization, and clearly deserves further attention. Also, since the energy released by solute precipitation adds to the energy released by grain growth, the combined mechanism of discontinuous precipitation provides a powerful energy source for microstructure change.

H. Phase transformation

Some of the metals used in semiconductor processing have structural phase transformations that may occur within the range of temperatures encountered in fabrication. These transformations are to be distinguished from the phase changes caused by compound formation. Examples include:

- (1) cobalt, which has a reversible first order transformation from hcp ϵ -Co to fcc α -Co at 442 °C;³⁵
- (2) tantalum, which has an irreversible transformation from the metastable tetragonal β -Ta phase to the stable bcc α -Ta phase at approximately 750 °C or lower;³⁶
- (3) titanium disilicide, which has an irreversible transformation from the metastable orthorhombic C49-TiSi₂ phase to the stable orthorhombic C54-TiSi₂ phase at approximately 800 °C,³⁷ and
- (4) tungsten disilicide, which has an irreversible transformation from the metastable hexagonal C40-WSi₂ phase to the stable tetragonal C11b-WSi₂ phase at approximately 750 °C.³⁸

The Co phase transformation is reversible in bulk Co, but does not always exhibit this property in thin films.³⁵ For this phase change, the bulk heat of transformation is 0.51 kJ/mol,³⁹ corresponding to 5.3 meV/atom. However, the transformation is endothermic upon heating, thus no energy is released to drive grain growth. In fact, in thin film studies of Co,³⁵ it was found that grain growth and stress relief occurred during heating before the onset of the phase transformation so that this transformation had little effect on the overall stress and microstructure. The sources of energy available in the irreversible transformations (Ta, TiSi₂, WSi₂) are considered small compared with heats of compound formation, for example, the free energy difference between the TiSi₂ phases is estimated to be no larger than 600 MJ/m³, or 50 meV/atom (Ref. 40) (Fig. 4). This value is consistent with a low density of nucleation sites for the transformation to the C54-TiSi₂ phase. However, it is higher than the energy densities involved in grain growth, interface energy minimization, and strain energy minimization, and is comparable to the energy density available in discontinuous

precipitation. As will be described, the transformation to C54-TiSi_2 can produce a strongly oriented microstructure in very narrow lines.

I. Energetic particle bombardment during deposition

Microstructure modification by particle bombardment during thin film deposition usually requires a bombardment energy on the order of 1–10 eV per deposited film atom (Fig. 4) to obtain measurable structure changes^{41–44} For Cu, this represents an energy input of approximately 100 GJ/m³ of film volume, most of which is dissipated as heat. Clearly particle bombardment seems a very inefficient way to modify the microstructure of a material. However, it may only be necessary to modify a thin layer of material, such as a refractory metal underlayer, in order to obtain desirable microstructure changes in a layered structure. Also, the directional quality of particle bombardment, which is not available in the other mechanisms listed above, may be exploited to provide an in-plane orientation of the film microstructure. For significant modification of in-plane orientation in refractory metal films, it has been found necessary to deliver more than 100 eV/atom to the growing film.^{45,46}

J. Mechanical deformation

Mechanical deformation of bulk metals is used to change not only their shapes but also their microstructures and mechanical properties. In soft metals such as Cu, the strain energy density added to the metal by 30% deformation (e.g., by cold working at $T < 0.5T_m$) is on the order of 25 J/mol,⁴⁷ equivalent to 3.5 MJ/m³ or 0.3 meV/atom. This energy may be increased significantly to 120 MJ/m³, or 10 meV/atom (Fig. 4), with further deformation, which can be achieved by adding alloying elements to the metal or by carrying out the deformation at low temperature where the damage does not anneal out spontaneously. Mechanical deformation stores energy in dislocation networks and their associated strain fields. In a soft bulk metal, dislocation densities of 10^{10} – $10^{12}/\text{m}^2$ are obtained with moderate levels of deformation.⁴⁷ With severe deformation, such as cold rolling, dislocation densities of $10^{16}/\text{m}^2$ are generated.¹⁸ The line energy in a dislocation is given by Gb^2 , where G is the shear modulus and b is the Burgers vector of the dislocation.¹⁸ If annealing eliminates a high density of dislocations (ρ dislocations/m²), replacing them with a recrystallized region of low dislocation density, the stored energy released is given by ρGb^2 . With a dislocation line energy of $Gb^2 = 10^{-8}$ J/m,¹⁸ and a dislocation density $\rho = 10^{16}/\text{m}^2$, the energy density is about 100 MJ/m³, or 8 meV/atom, which is comparable to the energy densities for grain growth and solute precipitation. This dislocation density corresponds to a dislocation spacing of about 10 nm, which is not, however, observed in Al–Cu metallization, where most grains appear to be free of dislocations.⁴⁸ The high purity of the materials, the constraints applied by the substrate, and the proximity of surfaces and interfaces prevent the accumulation of significant dislocation densities in semiconductor metallization. It is possible, however, that processes such as CMP may intro-

duce transient dislocation densities with sufficient energy density to affect the conductor microstructure. Whether this energy can be used for post-CMP microstructure control remains to be demonstrated. Physical deformation also occurs during processes such as the filling of vias or trenches using high pressure to extrude an Al–Cu film into the recessed features.²⁶ These filling processes, however, are usually carried out at elevated temperature ($\sim 400^\circ\text{C}$), which prevents a high dislocation density from remaining in the metal film. Therefore stored energy from physical deformation is not presently a significant mechanism to drive microstructure change in semiconductor metallization, although it could potentially be introduced into a fabrication process.

The energy stored in physical deformation is released by a variety of processes including recovery, grain growth, recrystallization, and secondary recrystallization,¹⁸ all of which modify the microstructure. Recovery refers to the annihilation of point defects and rearrangement of dislocations, grain growth (discussed earlier) refers to the elimination of grain boundary area, recrystallization refers to the migration of large angle grain boundaries within a deformed matrix, and secondary recrystallization (sometimes referred to as abnormal grain growth) occurs when the grain size is temporarily limited during annealing by some factor such as geometry, solute pinning, or the development of a strong preferred orientation.¹⁸ With continued heating, secondary recrystallization may start by a nucleation process, overcoming the effects of solutes and free surfaces, and continue until the matrix is completely transformed. Given the very small dimensions of future interconnection structures (Fig. 1), it is clear that complex dislocation networks will not be formed, and only a subset of these mechanisms will be available. It may, however, be possible to take advantage of highly oriented textures which are formed under mechanical deformation.¹⁸ One example where cold rolling was successfully used to control thin film microstructure is in the fabrication of template substrates for oriented growth of the high temperature superconductor yttrium–barium–copper–oxide (YBCO) and related compounds.⁴⁹ It was shown that rolling deformation of a nickel substrate introduces an in-plane orientation which controls the microstructure of the yttria-stabilized zirconia (YSZ) template layer deposited at high temperature ($\sim 600^\circ\text{C}$). This template layer subsequently provides an in-plane orientation to the YBCO superconductor film, resulting in improved superconducting properties.

K. Radiation damage

Ion implantation or other energetic particle irradiation introduces stored energy in the form of vacancies and interstitial atoms, and the accumulation of radiation damage is characterized by the number of displacements per atom (dpa).^{18,50} The displacement energy for Si atoms is 15 eV, and for intermetallic compounds such as NiAl about 25 eV.⁵⁰ Doses sufficient to displace atoms from their equilibrium lattice sites clearly cause severe atomic rearrangement, which can lead to large local defect gradients. Upon annealing,

some of these defects will be eliminated, since they are energetically unfavorable. Those that remain act like solute atoms, in addition to intentionally implanted species, and can impede dislocation motion and modify diffusion coefficients, thereby affecting microstructure evolution.⁵¹ An important advantage of ion implantation is the ability to control the magnitude of the flux, the depth of penetration, and the lateral location of the implant (by masking). Therefore, large gradients of solute and defect concentrations can be introduced into a film, allowing one to modify the microstructure on a local scale. As the dosage is increased to several tenths dpa, the damage reaches the threshold for amorphization of Si or metallic compounds such as NiAl,⁵⁰ in which case the stored energy becomes comparable to that obtained from crystallization or solidification, or higher (see Table I). This energy source is shown in Fig. 4 as a range including a higher energy portion (unshaded) to represent the thermal energy input which is not available for direct control of microstructure. A variation of introducing energy by ion implantation is to carry out the implantation at a temperature where grain growth can occur. Enhancement of grain growth by ion bombardment has been demonstrated for several ion and film species,⁵² and is found to depend directly on the defect yield per incident ion. The energy is delivered by elastic collisions at or near grain boundaries, and the resulting grain size is a function of the total ion dose. This method may be considered as a supplement to grain growth by thermal annealing in cases where the temperature is limited. Using ion implantation to generate stored energy in patterned damascene metallization requires additional blocking masks to avoid damaging the insulator. Also, the substantial doses required for microstructure change will lengthen the fabrication process. These factors make it unlikely that ion implantation damage of the conductor metals will be widely used to control microstructure in semiconductor metallization.

The sources of energy density discussed in Sec. II (and listed in Table I and Fig. 4) provide a wide range of mechanisms to influence thin film microstructure. In Secs. III and IV, examples of microstructure control during and after deposition are discussed in terms of these energy sources.

III. EXAMPLES OF MICROSTRUCTURE CONTROL DURING DEPOSITION

A. Effects of the underlying material on the texture of Ti/Al-Cu

1. Oxide surface roughness

Interface energy minimization has a strong orienting influence on thin films when the substrate surface is flat, but when the surface is rough, the orientations of nearby regions differ, and the overall texture may be weak. For example, the quality of the SiO₂ surface underlying Ti/Al-Cu/Ti metallization has been shown to directly affect the microstructure of Al-Cu.⁵³ In a recent study by Rodbell *et al.*,⁵⁴ the effects of surface roughness on the Ti layer and the Al-Cu layer were measured separately. A smooth SiO₂, with a root-mean-square (rms) surface roughness of 0.6 nm formed by plasma-assisted tetraethylorthosilicate (TEOS) oxide deposition, was

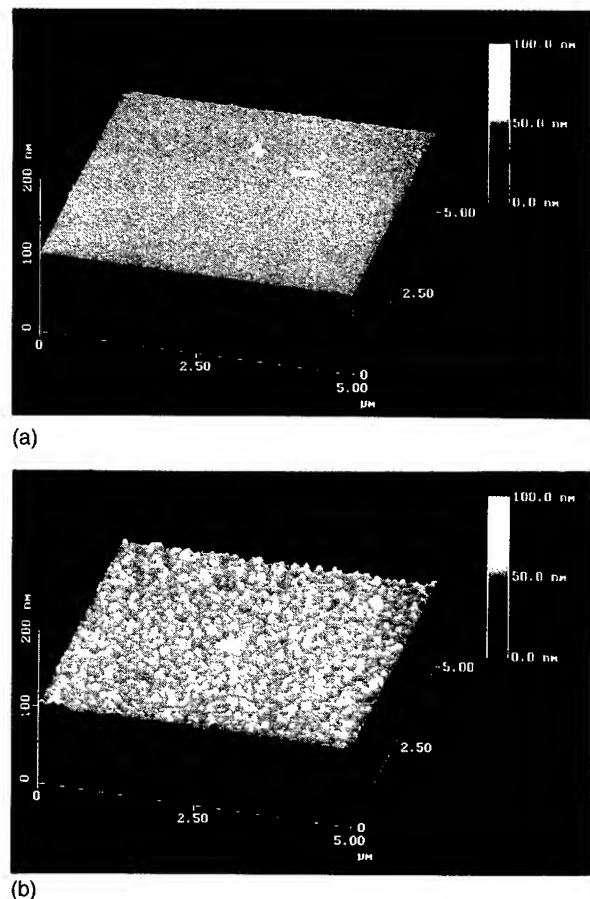


FIG. 6. Comparison of surface roughness measured by atomic force microscopy for (a) smooth and (b) rough SiO₂ coated with 20 nm Ti (Ref. 55).

compared with a rough SiO₂ with a rms surface roughness of 2.8 nm formed by plasma-assisted silane oxide deposition. These roughness values were obtained for an image size of 5×5 μm. The roughness of these two oxides is evident in Fig. 6, which shows atomic force microscopy (AFM) images of the two surfaces coated with 20 nm Ti at 150 °C, prior to Al-Cu deposition.⁵⁵ The texture of the 20 nm Ti layer is difficult to measure at this thickness, so a 200-nm-thick Ti layer was deposited onto the two SiO₂ surfaces and measured using the Schulz x-ray pole figure method. This method uses an x-ray diffractometer set for a particular Bragg reflection [e.g., (0002) for Ti], and maps the intensity of this reflection on the hemisphere above the sample surface. The (0002) pole figures for the smooth and the rough oxides are shown in Figs. 7(a) and 7(b), respectively. The Ti films on both oxides were found to be oriented with a (0002) fiber texture, indicating that the hexagonal basal planes were parallel to the plane of the substrate. The strength of this fiber texture, measured by the angular width of the (0002) peak, was substantially higher on the smooth oxide (5° peak width) than on the rough oxide (18° peak width). Ti deposited by collimated sputtering, which decreases the angular spread of incident atoms, also has a stronger fiber texture than Ti deposited by conventional magnetron sputtering.⁵⁴

For 500 nm Al-Cu deposited directly onto the two SiO₂

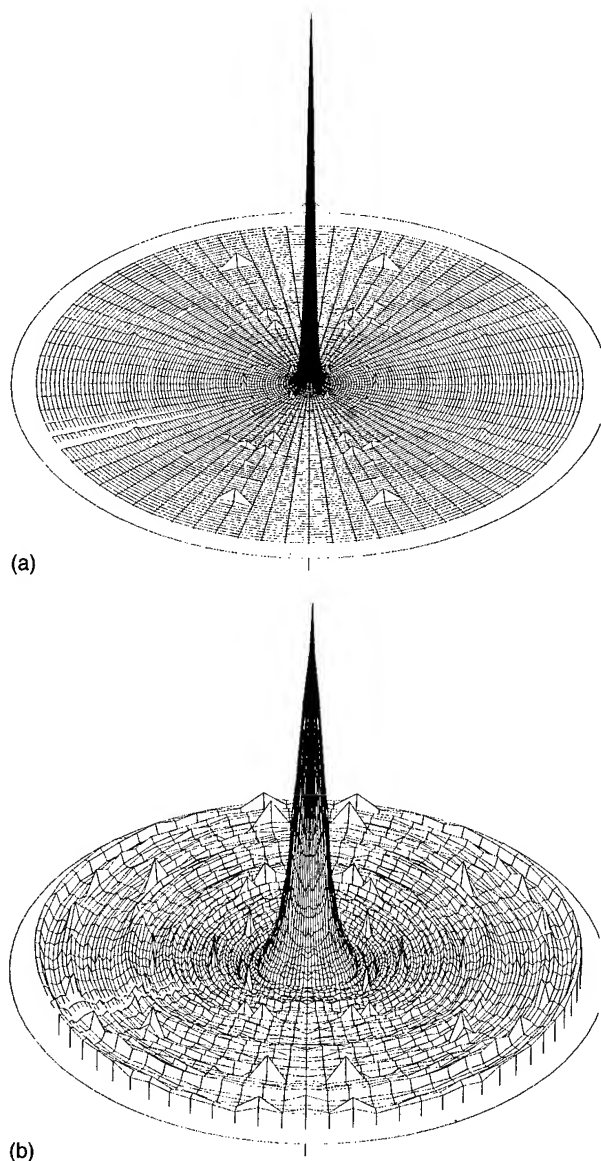


FIG. 7. Comparison of (0002) x-ray pole figures for 200-nm-thick Ti films deposited on (a) smooth and (b) rough SiO_2 (Ref. 54).

surfaces, the pole figures [Fig. 8(a)] showed a near-(111) fiber texture, in which the orientation with the highest (111) intensity is not normal to the substrate plane, but lies on a circle offset from the normal. Since these pole figures are azimuthally symmetric, they are plotted in Fig. 8 only as a function of tilt angle. For the smooth oxide surface, the offset angle was found to be 3° , and for the rougher oxide it was 4.5° . The angular width was 12° on the smooth oxide and 16° on the rough oxide. This difference is also evident in the angular widths of the (111) peaks located at 70.5° [Fig. 8(a)]. The volume fraction of randomly oriented grains [Fig. 8(a)], was also lower on the smooth oxide (4%) than on the rough oxide (8%). For these oxide substrates, it appears that the range of local surface orientations affects the textures of both Ti and Al-Cu, although these two materials show different responses. Predicting the effects of surface roughness on thin film texture will require taking into account interface energy

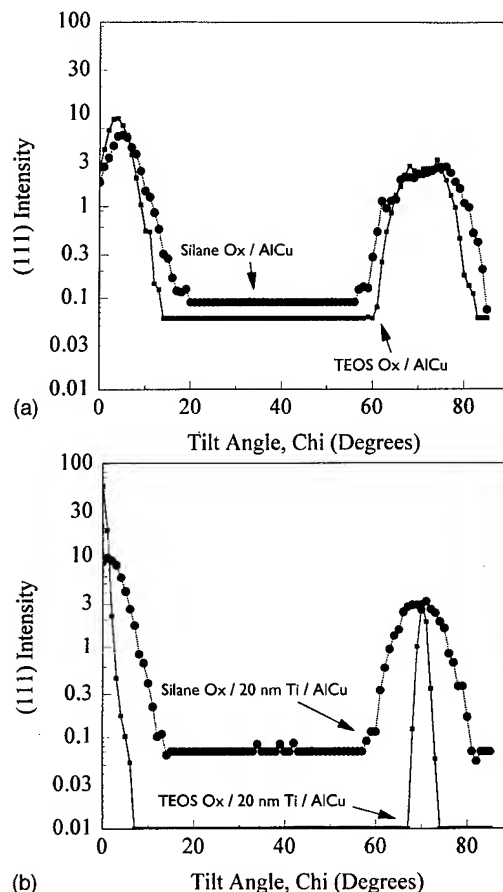


FIG. 8. Comparison of (111) x-ray pole figures for 500-nm-thick Al-Cu films deposited on (a) smooth vs rough SiO_2 , and on (b) 20-nm-thick Ti on smooth vs rough SiO_2 (Ref. 54).

minimization together with the coalescence of islands with a range of orientations, followed by grain growth during deposition to produce the measured texture. This is an area where more quantitative analysis will provide better methods for microstructure control. In other studies, the microstructure of Cu was also shown to change from a porous, columnar morphology on rough Al_2O_3 substrates to a dense, noncolumnar morphology on smooth Al_2O_3 substrates.⁵⁶ A determining factor is the ratio of the adatom diffusion length to the characteristic length of the surface roughness. A strongly oriented, dense film is favored when the adatom diffusion length exceeds the length scale of the surface roughness. Under these conditions, the deposited film atoms have enough surface mobility to establish an orientation with a low interface energy with respect to the substrate. If the substrate is too rough, however, interface energy minimization cannot overcome the range of orientations imposed by the substrate surface angles.

2. Ti underlayer texture

When an Al-Cu film is deposited onto a textured Ti layer, as in the Ti deposited onto SiO_2 surfaces described earlier, the Al-Cu demonstrates "texture inheritance,"⁵⁴ gaining a stronger (111) texture than that obtained on the oxide surface

alone. The difference between Ti/Al-Cu on a smooth (TEOS) oxide and on a rough (silane) oxide is illustrated in Fig. 8(b). The 500-nm-thick Al-Cu film on 20 nm Ti on the smooth oxide has a pure (111) fiber texture with an angular width of 6° and a low random fraction (1%), compared to the near-(111) texture and higher random fraction (4%) observed on the smooth oxide alone, without the Ti underlayer [Fig. 8(a)]. However, the Al-Cu on Ti on the rough oxide [Fig. 8(b)] shows a near-(111) texture, with an offset angle of 2° and an angular width of 14° , comparable to the Al-Cu on SiO₂ [Fig. 8(a)], which had an offset angle of 4.5° and width of 16° . These texture measurements demonstrate that the Ti underlayer has a dramatic effect only on the Al-Cu texture on the smoother oxide, as it cannot overcome the wide range of orientations on the rougher oxide. In addition to the mechanism of interface energy minimization, which orients the Ti with (0002) planes parallel to the oxide surface and the Al-Cu with (111) planes parallel to the oxide surface, the Al-Cu growing on highly oriented Ti also experiences a component of granular epitaxy with respect to the Ti grains, in which the Al-Cu grains inherit the orientation of the underlying Ti grains. Subsequently, Al-Cu grain growth during deposition overshadows granular epitaxy, since the Al-Cu grains readily grow much larger than the Ti grains. The orienting effect of the Ti layer, however, continues to influence the Al-Cu texture as the Al-Cu grains thicken and maintain their strong (111) texture. If the deposition temperature is high enough, Al₃Ti compound formation also influences the texture of the Al-Cu film, since the reaction product itself is strongly oriented with respect to the initial Ti layer.³²

B. Effects of geometrical confinement on Al-Cu texture

1. Texture of blanket patterned Al-Cu lines

When Al-Cu films are deposited as blanket films, then patterned by subtractive etching into linewidths of less than $1.0\text{ }\mu\text{m}$ and aspect ratios near unity, a bamboo microstructure develops,¹⁵ and the strength of the (111) texture is enhanced relative to blanket films or lines with aspect ratios less than unity.⁸ Also, in these narrow lines, the misorientation angles between grains are no longer randomly distributed. With a strong (111) texture, the grain boundaries are primarily tilt boundaries, with their grain-to-grain misorientation angles grouped around certain angles that correspond to low energy grain boundaries.⁵⁷ As the grain orientations are mapped along the line, it is also observed that the orientations occur in clusters, suggesting some further interactions during deposition or annealing.⁷

2. Texture of damascene processed Al-Cu lines

The mechanisms that determine grain orientations in patterned blanket Al-Cu differ from those in damascene Al-Cu processing, in which the metal is deposited into existing trench and via structures. An example is damascene processing of Al-0.5 wt % Cu using a "cold-hot" magnetron deposition sequence with a wafer temperature of about 400°C ,⁵⁸ in which the strength of the (111) texture was found to de-

crease as the linewidth decreased from 16 to $0.5\text{ }\mu\text{m}$, with a constant trench depth of $1.0\text{ }\mu\text{m}$. In the higher aspect ratio features, the sidewall interfaces exert more influence on the grain orientations than does the bottom interface. The main orienting mechanism is still interface energy minimization, but it has the greatest effect at the sidewalls in these structures.

C. Effects of energetic particle bombardment during growth on thin film texture

1. Bombardment modification of fiber texture

Sputter deposition of thin films supplies ion and neutral atom bombardment with energies ranging from thermal to hundreds of eV.⁵⁹ When the sputtering plasma boundary is approximately parallel to the plane of the substrate, as is the case in the typical magnetron deposition system, the incident direction of energetic particles is primarily normal to the surface, and therefore bombardment effects may be indistinguishable from the orienting effect of the substrate plane itself. Studies of texture effects under normal incidence particle bombardment have often shown a strengthening of fiber texture or in some cases a change of orientation of the fiber axis with increasing target power or decreasing gas pressure, while maintaining azimuthal symmetry in the plane of the substrate.^{60,61} The enhanced fiber texture is usually attributed to enhanced adatom mobility, somewhat analogous to a higher deposition temperature, and changes in fiber texture orientation are usually attributed to the resputtering of non-channeled crystallite orientations. For example, the preferred orientation of TiN on oxidized Si has been controllably changed from purely (111) to purely (002) by increasing the ion/neutral ratio from 1 to greater than 5 during 20 eV nitrogen ion bombardment of the growing film.⁶² By using low energy ion sources, it is possible to distinguish between the effects of gas ions and energetic film atoms or ions. An example is the ultrahigh vacuum (UHV) growth of Al films under controlled self-ion bombardment from an Al⁺ ion source, in which the ion energy and ion/neutral ratio can be independently controlled.⁶³ The full width at half-maximum of the Al (111) rocking curve was found to decrease from 9.6° with Al⁺ ion energy of 10 eV to 2.2° with Al⁺ ion energy of 120 eV for a fixed ion/neutral ratio of 0.68. It was also shown that the strength of the (111) texture was essentially determined during the nucleation and coalescence stages, and that granular epitaxy maintained the strong texture of this initial layer as the film thickness increased to 300 nm.

2. In-plane orientation caused by off-normal incidence particle bombardment

In the above examples, the energetic particle bombardment was directed primarily normal to the substrate plane, influencing the fiber texture while maintaining azimuthal symmetry. By supplying particle bombardment at off-normal incidence, additional alignment effects have been demonstrated. Low energy (100–300 eV) Ar⁺ ion beam bombardment at 70° from normal incidence during Nb film deposition

was shown to strongly align the bcc Nb grains relative to the beam direction, producing an in-plane azimuthal orientation.⁴⁵ In this experiment, the Nb was deposited by ion beam sputtering, and the depositing atoms had enough energy to develop a strong (110) fiber texture on amorphous silica substrates in the absence of bombardment. With the addition of off-normal ion bombardment, the (110) preferred orientation remained, while an in-plane orientation increased strength with increasing ion/atom ratio. It was shown^{64,65} that the direction selected for growth corresponded to the ion beam aligned with a planar channeling direction between sets of (110) planes in the grains, which thereby experienced less removal by resputtering than other grain orientations. A similar approach has proven successful in orienting films of transition metal oxides, such as yttria-stabilized ZrO_2 (YSZ), which is used as a template for the oriented growth of the high-temperature superconductor YBCO.⁶⁶ YSZ has a cubic structure and grows with a strong (001) fiber texture on amorphous or polycrystalline substrates at elevated temperature. Ion bombardment has been applied at angles corresponding to the (111) axis⁶⁶ as well as the (110) axis,⁶⁷ with strong in-plane orientations achieved. Since YSZ is an ionic compound, an additional mechanism has been proposed to strengthen in-plane orientation, in which growth rate anisotropy is related to the requirement for charge neutrality on the growing crystal faces.⁶⁷ Other bombardment effects reported in this material system⁶⁸ include tilting of the YSZ columnar microstructure towards the ion beam direction and clustering of neighboring grain orientations similar to the fine-line Al-Cu described above.⁷

Most sputter deposition chambers for silicon wafer processing are cylindrically symmetrical around the axis of a circular wafer, and therefore do not have a geometry that provides off-normal incidence particle bombardment, except for the difference in symmetry between the center and edges of the wafer. A more pronounced asymmetry is present in linear scanning sputtering systems, such as those used for flat panel displays and large area optical coatings. In these systems, the substrate enters the deposition region from one side, passes in front of the sputtering target where the highest deposition rate occurs, and exits from the opposite side. Asymmetric properties including stress⁶⁹ have been attributed to the different angular distributions of arriving species. The distributions of energy and direction of arriving atoms and ions in the plane which includes the wafer normal and the scan direction differ from the distributions in the plane which includes the wafer normal and the perpendicular to the scan direction. An example of such an asymmetric microstructure in sputter deposited bcc Mo is given in Fig. 9, which shows the (110) x-ray pole figure for a film deposited on oxidized Si in a linear scanning sputtering chamber.⁷⁰ As with the bcc Nb films described above, the Mo film deposition process is energetic enough in this low pressure sputtering environment to form a strong (110) fiber texture on the amorphous substrate, shown by the strong central peak of the (110) pole figure. Also evident, however, is the azimuthal asymmetry of the intensity of the ring at 60° , which corre-

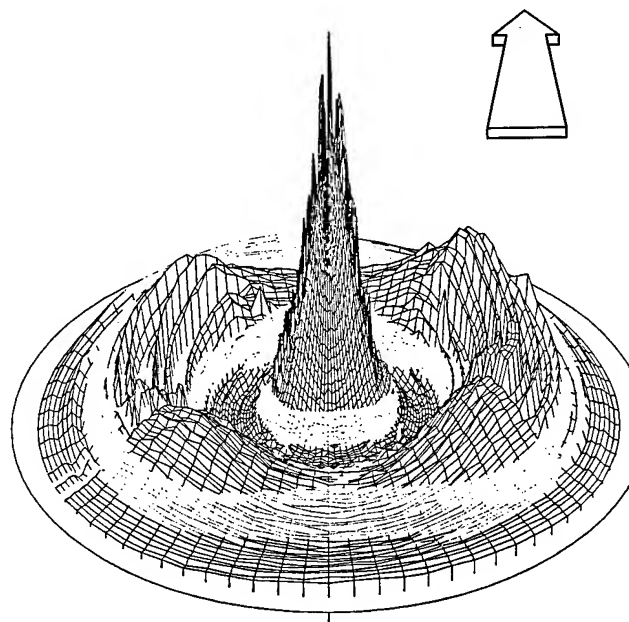


FIG. 9. (110) x-ray pole figure for a 500-nm-thick Mo film deposited on oxidized Si in a linear scanning sputter deposition system. The direction of the substrate motion relative to the target during deposition is shown by the arrow (Ref. 70).

sponds to (110) planes tilted with respect to the substrate normal. The azimuthal positions of the four higher intensity peaks in the (110) ring at 60° show the relationship of the oriented microstructure to the scan direction, which is the front-to-back direction in Fig. 9.

An in-plane orientation in Mo films was also obtained by placing samples off the axis of a circular target sputtering system or at a tilted angle with respect to the sputtering target plane.⁷¹ In-plane orientation in Mo(N) films was also observed in a stationary rectangular target magnetron sputtering system, in which the different arrival angles from the long and short directions of the target were shown to induce orientation.⁷² The resulting orientations relative to the direction of bombardment or deposition depend on the specific angles and flux ratios occurring during film growth. As mentioned above, inert ion beam bombardment during deposition provides a controllable process for modifying the bombardment angles, energies, and flux values without changing the deposition flux parameters.⁴⁵ Using argon ion beam bombardment during Nb deposition,⁴⁶ it was shown that bombardment at an angle near 45° from normal induces a restricted (110) fiber texture in which the ion beam aligns with the (200) direction at 45° to the substrate, and bombardment near 35° from normal induces a restricted (110) fiber texture in which the beam aligns with the (111) direction. At least three types of in-plane orientation with (110) fiber texture have been documented in bcc films of Mo and Nb under bombardment at various angles and flux conditions.^{45,46} It was also found that, when the thickness of Mo is increased above about $1\ \mu\text{m}$, strain energy effects begin to influence the film orientation, and a change from strong (110) fiber

texture to strong (111) fiber texture was reported.⁷¹

It is quite interesting that a clear azimuthal asymmetry often emerges from a scanning substrate motion, since the angles of incidence of atoms and ions are not constant during the scan. This observation is consistent with the crystallographic orientation being set at an early stage of film growth, even before the substrate enters the main deposition region. It is therefore clear that some materials may be strongly oriented both normal to the substrate plane and within the substrate plane by energetic particle bombardment at the appropriate stage of film growth.

One more recent example of in-plane orientation shows that deposition at almost glancing angle (75° from normal incidence) may induce strong in-plane orientation when it is coupled to a strong fiber texture mechanism. This appears to be the case in the deposition of AlN by sputtering at various angles onto glass substrates,⁷³ without added ion bombardment. AlN films deposited at normal incidence showed the usual strong (0002) fiber texture, in which the hexagonal base plane of the wurtzite structure is parallel to the substrate plane. With increasing tilt angle, the microstructure developed two populations of oriented grains and, when the deposition angle was near glancing (75° from normal), the texture became sharp in both the normal direction and the in-plane direction, producing a highly aligned microstructure. The selection of the in-plane orientation is consistent with channeling of the energetic depositing atoms and reflected sputtering gas atoms along directions of low sputtering yield, allowing enhanced growth of the oriented grains relative to nonchanneled orientations. This result adds to the evidence that in-plane highly oriented microstructures may be obtained in some materials by simple deposition geometries.

IV. EXAMPLES OF MICROSTRUCTURE CONTROL AFTER DEPOSITION

A. Grain growth coupled to solute precipitation

As described earlier, there is a significant amount of energy released when grain growth is coupled to solute precipitation in the process of discontinuous precipitation. An example of a dramatic texture change caused by discontinuous precipitation is found in thin films of Cu-Co alloys.¹³ With concentrations of 0.4–8.6 at. %Co in coevaporated or electroplated Cu-Co thin films, the as-deposited films showed a strong (111) fiber texture (Fig. 10). After annealing at 450 °C for 120 min, however, the texture had changed to a strong (200) fiber texture (Fig. 10), with many individual grains larger than 10 times the film thickness. This behavior is not observed in pure Cu films. Measurements of stress, resistivity, and grain structure [by transmission electron microscopy (TEM)] as a continuous function of temperature showed that the precipitation of supersaturated Co and significant grain growth both occurred at about 200 °C, and that this temperature of precipitation increased with Co concentration. In the as-deposited coevaporated films, magnetic measurements indicated that about 65% of the Co atoms were already precipitated as particles about 9 nm in diameter before annealing. Therefore, the energy from precipitation available to

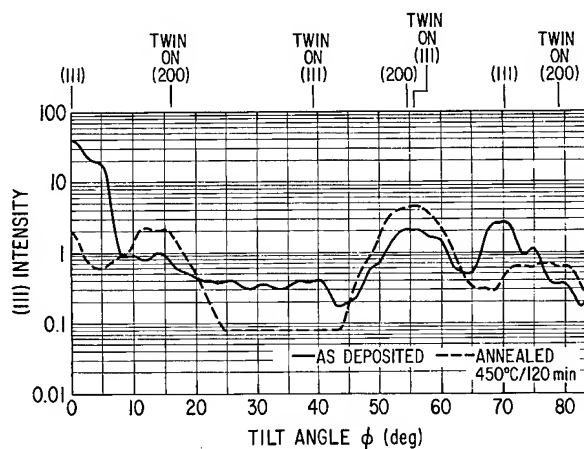


FIG. 10. (111) x-ray pole figure of a sputtered Cu-0.44 at. % Co film before and after annealing at 450 °C for 120 min (Ref. 13).

drive grain growth was reduced to about one-third of that available if all the Co was still in solution in the as-deposited film. It was clear, however, that the growing grains had a high enough driving force to overcome being pinned by the precipitated Co particles.

The balance between grain boundary mobility and grain boundary pinning during discontinuous precipitation is a delicate one, and it has been shown in bulk Cu-Co alloys that the processes of grain growth and precipitation can occur in either order, depending on the solute concentration and the presence of additional stored energy from deformation.⁷⁴ In studies of bulk Cu-2 wt %Co alloys,⁷⁵ the pressure behind pinned grain boundaries (equivalent to energy density) was measured directly from the curvature of the boundaries and the spacing of the pinning precipitate particles. For a Co particle spacing of 170 nm, which corresponded to an average particle diameter of 13 nm, the pressure was found to be approximately 7 MPa,⁷⁵ equivalent to an energy density of 7 MJ/m³, or 0.5 meV/atom. In this work, the Co particles were already present in the highly deformed material as the grain boundaries moved from the less deformed material, so that the shapes were measured when the driving force for grain growth and the pinning force were in balance. Comparing this with the case of Cu-Co discontinuous precipitation described above,¹³ the energy density available from solute precipitation is about two orders of magnitude higher than the grain boundary pinning energy density, which is consistent with the observed presence of Co particles left behind in the interior of large Cu grains rather than located only at Cu grain boundaries.¹³ Pinning of grain growth has been observed in layered Al thin films,⁷⁶ for which normal grain growth is inhibited at low temperature. As the temperature is increased, thermal mismatch strain and increased grain boundary mobility eventually allow grain boundaries to escape the pinning force of the interfaces in the layered microstructure.⁷⁷ At this point, abnormal grain growth can occur throughout the film volume, accompanied by changes in texture.⁷⁶

A related example is found in the microstructural evolu-

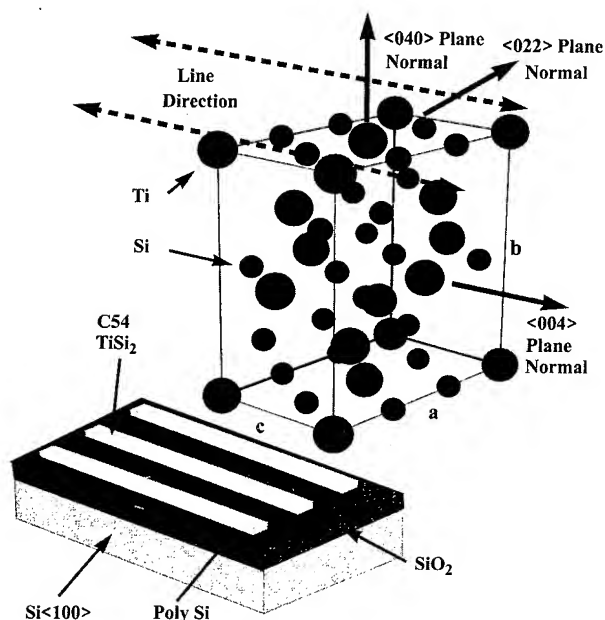


FIG. 11. Schematic diagram of the crystal orientation of C54 TiSi₂ relative to the line direction in submicrometer arrays indicating the orientations of the (040), (004), and (022) planes normal (Ref. 81).

tion of codeposited Cu_{0.3}Mo_{0.7} films,⁷⁸ which form a nanocrystalline single phase bcc solution as deposited, and phase separate on annealing at 450 °C into fcc Cu and bcc Mo, which are normally mutually insoluble. The energy released by phase separation facilitates Mo grain growth and an increase in (110) texture, neither of which occurs in pure Mo at this temperature. These examples suggest that even more pronounced microstructure changes might be possible with the right choice of low-solubility solute species introduced with a deposition process that maintains a high supersaturated solute concentration until annealed above the deposition temperature.

B. Grain growth coupled to phase transformation in narrow lines

Titanium disilicide (TiSi₂) occurs in two crystal structures, the C49 phase and the C54 phase. In the processing of microelectronic silicide contacts by reaction of Ti with Si, the C49 phase appears first upon heating at about 700 °C.³⁸ Since this phase has a relatively high resistivity (60–75 μΩ cm), it is necessary to heat it to around 800 °C to obtain the C54 phase with a low resistivity (15–20 μΩ cm). As mentioned earlier, the free energy difference between the two phases, estimated to be less than 600 MJ/m³,⁴⁰ is relatively small compared with heats of reaction, so that the transformation has a low density of nucleation sites within the C49 phase. This low nucleation density means that even though the C49 TiSi₂ phase forms readily wherever Ti is in contact with Si, there is very little driving force to transform the C49 phase into the C54 phase. Therefore, when the transformation does occur in narrow patterned lines, it has been observed to propagate from widely spaced nucleation sites,

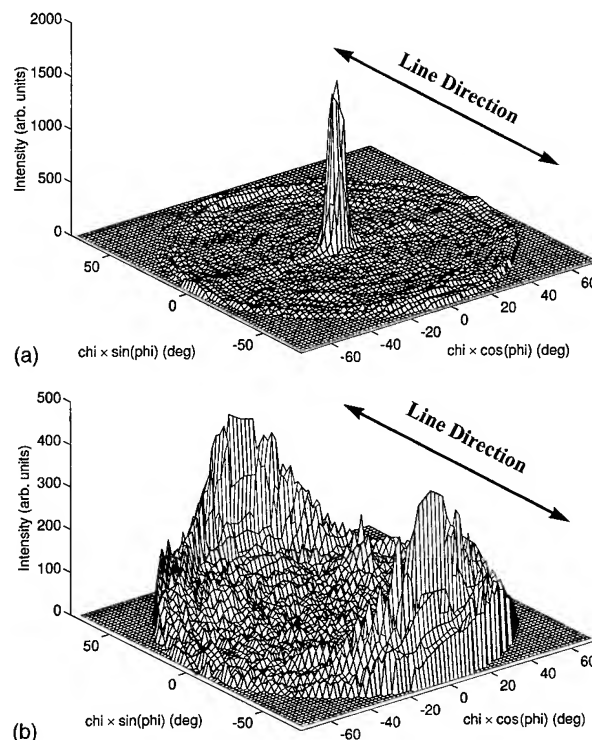


FIG. 12. (a) (040) and (b) (022) x-ray pole figures of C54 TiSi₂ after transformation from the C49 TiSi₂ phase, in an array of 0.2-μm-wide lines. The horizontal axis labels indicate the tilt angles relative to the substrate normal (Ref. 81).

and move along the silicide line converting fine grained C49 to large grained C54 TiSi₂.⁷⁹ The grain size in the C49 phase is typically tens of nanometers, and in the C54 phase it is typically many micrometers, substantially larger than the thickness and width of the line. Several sources of energy are available to drive this microstructure change, namely, energy released by the transformation itself, interface energy minimization, and the energy released by eliminating the high density of C49 grain boundaries. This grain boundary energy is released by the process of phase transformation, and may be distinguished from the grain boundary energy of the growing C54 grains. The C54 grain boundary energy is not released unless competitive grain growth occurs within the C54 phase after the impingement of the newly formed grains. X-ray analysis shows that the C54 phase formed in narrow submicrometer lines develops both a strong out-of-plane texture and a strong in-plane texture related to the patterned direction of the narrow lines.^{80,81} The orientation of the grains (Fig. 11) shows the (040) plane normal oriented perpendicular to the substrate surface, and the (004) plane normal oriented parallel to the patterned line direction. The (040) pole figure is shown in Fig. 12(a) for C54 TiSi₂ lines of 0.2 μm width, indicating a strong (040) fiber texture.^{80,81} It is impractical to measure the orientation of the (004) planes directly, since their normal is in the substrate plane. Instead, the azimuthal orientation was determined by measuring the pole figure of the (022) planes, which form an angle of 60.7° with the (040) planes. This pole figure, shown in Fig. 12(b),

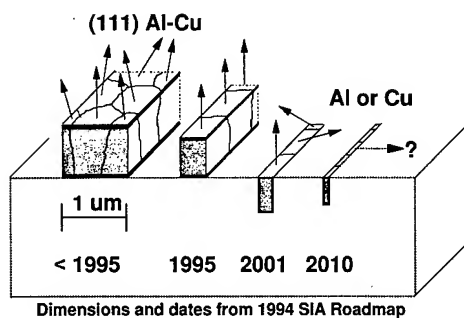


FIG. 13. Schematic diagram of the trends in microstructure of semiconductor metallization, based on dimensions and dates from the 1994 SIA Roadmap (Ref. 1). The arrows indicate (111) orientations of the Al-Cu or Cu grains.

has distinct intensity maxima in the direction of the patterned lines.⁸¹ These orientation measurements confirm that the C49-C54 phase transformation caused a major change in microstructure including a strong alignment of the grains relative to both the substrate plane and the line direction. The evidence suggests that the C54 grain orientation with the highest growth rate, apparently the (004) direction, is able to propagate along the line when oriented in this direction, while slower-growing orientations are overshadowed and therefore do not appear strongly in the x-ray measurement. The significance of these measurements is that the microstructure change stimulated by the phase transformation triggers an orientation selection mechanism, or combination of mechanisms, which produces a new microstructure with a highly oriented configuration.

V. PROSPECTS FOR CONTROL OF MICROSTRUCTURE IN SEMICONDUCTOR METALLIZATION

Having discussed the sources of energy available to drive thin film microstructure changes, and illustrating them with examples, we can now assess the prospects for increased control of microstructure in future semiconductor metallization. The evolution of conductor dimensions up to the present state of technology¹ has been to decrease the linewidth without decreasing the thickness proportionally, thereby increasing the aspect ratio to about 1.5:1, as shown in Fig. 1. The microstructure of Al-Cu lines, shown schematically in Fig. 13, has evolved from polycrystalline in wide lines to near bamboo in lines of about 0.4–0.5 μm in width. Since these lines are fabricated by blanket deposition, usually with Ti underlayers, their texture is strongly (111), as indicated by the arrows in Fig. 13. As we look to the future, however, not only do the conductor dimensions continue to decrease, but the aspect ratio increases to values of around 4:1, and the lines will be fabricated more frequently by damascene processes in patterned insulators. There is evidence already^{28,58} that shows that the strong (111) fiber texture is weakened in such features, as can be seen in Fig. 13, because the sidewalls exert more influence than the bottom interface of the conductor. If we are to maintain control of the conductor microstructure in future technology, we must

identify helpful trends in very narrow conductors that will enable us to control the application of the energy sources discussed above. Several of these trends are discussed next.

A. Decreasing diffusion distances and numbers of atoms

As the linewidth of semiconductor metallization decreases to less than one thousand metal atoms, one clear trend is the decreasing distance between interior atoms and the nearest interface or material boundary. Since the diffusion distance is given by $2(Dt)^{1/2}$, where D is the diffusivity and t is the time, the time needed for atoms to diffuse from the interior to the boundary of an interconnection wire decreases quadratically with linewidth. For metals, values of the diffusivity D at a temperature of half the absolute melting temperature ($T_m/2$) are on the order of 10^{-15} cm^2/s for lattice diffusion, and 10^{-9} cm^2/s for grain boundary diffusion.^{82,83} Annealing at this temperature for 30 min gives diffusion lengths of about 20 nm and 20 μm , respectively, for lattice and grain boundary diffusion. Depending on the grain structure, this means that annealing the conductor metal at half the absolute melting temperature (400 °C for Cu) for short times (30 min) provides sufficient time for interior solute atoms to diffuse along grain boundaries and to come into equilibrium with the boundary interfaces of the conductor. If grain growth occurs, moving grain boundaries provide fast diffusion paths for interior solute atoms to redistribute to the boundary interfaces, providing a sink for solute atoms with low solubility. A consequence of decreasing dimensions is the small number of atoms. For example, a 0.1- μm -wide copper line is only about 400 atoms wide. The details of microstructure changes, especially the energetics of nucleation and propagation, differ significantly in such narrow lines from the mechanisms in bulk materials or blanket thin films due to the small number of atoms and the proximity of interfaces. If a Cu line 0.1 μm wide contains 5 at. % solute atoms, the linewidth contains only about 20 solute atoms. In this case, when solute precipitation occurs, it is unlikely that nucleation of a precipitate particle can occur within the interior of the line. More likely, the sites for precipitation will be the outer boundaries of the line at interfaces with the surrounding material. If segregation of the precipitated solute atoms is strong, as it is when the solubility is low, then solute atoms will accumulate and form a layer adjacent to the interfaces surrounding the Cu line. This process has the potential of improving the diffusion barrier properties of the surrounding interfaces, and will reduce the resistivity of the Cu conductor line. Additionally, if solute atoms move to the interface boundaries, the benefits of solute precipitation in driving microstructure change may be achieved without encountering particle pinning of the moving grain boundaries.

B. Decreasing density of nucleation sites

A second trend in very narrow patterned structures is the decrease in the density of nucleation sites for specific microstructure changes. An example is the C49-C54 TiSi_2 phase transformation, which has been shown to be sparsely nucle-

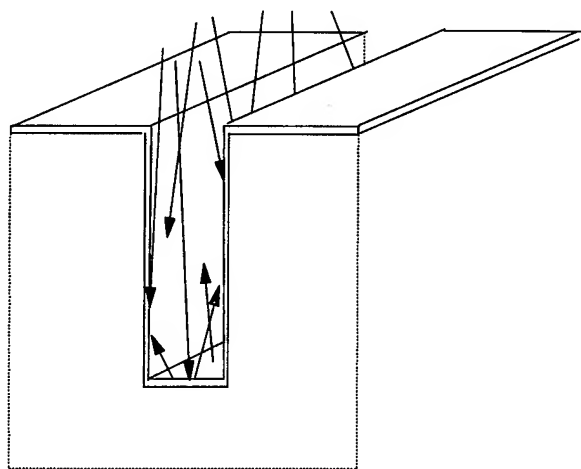


FIG. 14. Sketch of energetic ions and atoms bombarding the sidewalls in a damascene trench.

ated in narrow lines below a $1\ \mu\text{m}$ width. Although the atomic configuration of the nucleation sites for this transformation has not been identified, it was shown that nucleation is initiated at a small fraction ($\sim 15\%$) of grain boundary triple junctions in the C49 phase.⁸⁴ As linewidths decrease into the sub- $0.5\ \mu\text{m}$ regime, the density of nucleation sites with the correct atomic configuration decreases, with the result that some lines fail to transform completely, even with extended annealing.⁸⁵ A similar trend is expected for other microstructure changes where specific nucleation sites are required, such as in grain growth. As a result, we can expect to see microstructure changes becoming more sparsely nucleated in very narrow lines. This trend opens the possibility of more precise control of the nucleation of new grains and their propagation along the line to obtain a more complete modification into a new structure.

C. Increasing importance of sidewalls

As the aspect ratio of the conductor lines in semiconductor metallization continues to increase above 1:1, the tendency for the grain orientations to be determined by the sidewalls instead of the substrate plane suggests that it might become harder to generate strongly oriented conductor microstructures. However, one emerging possibility is that the microstructure of the barrier/seed layer in high aspect ratio features might be modified by ion bombardment (or self-bombardment) which is near-normal to the substrate and near-glancing incidence to the sidewalls (Fig. 14). This effect may produce an oriented seed layer on which the main Al or Cu conductor is deposited by physical vapor deposition (PVD) or non-PVD methods, and from which this layer derives its orientation. The combination of low nucleation site density and sufficient stored energy (e.g., from supersaturated solute precipitation) may provide the combination of mechanisms needed to grow strongly oriented conductors, if desired, with tightly controlled microstructures and low resistivities.

VI. CONCLUSIONS AND OUTSTANDING ISSUES

The microstructure of semiconductor metallization develops both during and after deposition under the influence of a combination of energy sources. The magnitude of these energy sources (Table I and Fig. 4) ranges from values of less than $1\ \text{meV/atom}$ that are involved in grain growth during thermal annealing to values exceeding $100\ \text{eV/atom}$ involved in energetic particle bombardment during deposition. Among the mechanisms described in this article, discontinuous precipitation stands out as a source of significant stored energy density that could be harnessed to gain further control over microstructure evolution in future metallization. This mechanism, which combines precipitation of supersaturated solute atoms with simultaneous grain growth, provides a continuous source of energy per unit volume of conductor metal as a grain boundary advances through the metal. The decreasing dimensions of metallization guarantee that interfaces and surfaces are close to all the interior atoms of the conductor, introducing the possibility of precipitating solute atoms almost entirely at interfaces during a controlled annealing treatment. This result achieves the lowest resistivity of the conductor, and may improve the adhesion and diffusion barrier characteristics of the interfaces if the solute species is properly chosen. It is clear that the strong (111) fiber texture presently obtained in near-bamboo Al-Cu metallization will not be characteristic of damascene metallization of either Al-Cu or Cu, due to the requirement of depositing the metal film into patterned trenches and vias. Microstructure control in metallization might, therefore, deteriorate in the near future. On the other hand, since these metals typically require underlayers or liners for successful integration, it is possible that the deposition of the underlayer or liner metal by an energetic particle process, such as sputtering, might induce an oriented microstructure that subsequently controls the microstructure of the conductor metal. This possibility emerges because of the increasing importance of the trench sidewalls in nucleating grain orientations in high aspect ratio features. Examples from several material systems, including metals, oxides, and nitrides, demonstrate that energetic particle bombardment during deposition may be used to generate highly oriented textures both normal to and parallel to the substrate surface. It is unlikely, however, that such an orienting mechanism will be available in a purely thermal chemical vapor deposition process.

There are many as yet unresolved issues raised by considering microstructure control in future semiconductor metallization. One continuing challenge is to determine how much microstructure control is needed in order to obtain the required levels of performance and reliability. Judging from recent correlations of electromigration lifetime with microstructure details, it is likely that continued control of microstructure will produce steady gains in performance reliability, and manufacturing yields, but greater quantification is needed. The relative magnitudes of the stored energy sources shown in Table I and Fig. 4 need further study in order to engineer the release of stored energy at the appropriate point in the device fabrication sequence. The mechanism of solute

precipitation, in particular, needs further quantification to identify the best solute species, concentrations, and temperature sequences needed to benefit from this stored energy. Also, coupling of solute precipitation to grain growth and orientation selection has not been studied in patterned thin film structures, where submicrometer dimensions have great influence. Since most metallization systems will combine underlayers or liners with the conductor metals, the effect of oriented templates on thin film microstructure deserves more attention. Finally, since the trend in future semiconductor metallization is towards use of the damascene process, it will be necessary to better understand the effects of geometrical confinement on grain nucleation, growth, and orientation selection. There is no shortage of challenging, interesting issues to resolve in controlling the microstructure of future semiconductor metallization.

ACKNOWLEDGMENTS

The authors thank D. B. Knorr, D. A. Smith, E. G. Colgan, L. A. Clevenger, C. Cabral, Jr., R. A. Roy, V. Svilan, S. M. Rossnagel, C. Noyan, R. Rosenberg, and R. H. Hammond for collaborations and numerous discussions on thin film microstructure. They thank C. Lavoie and R. M. Geffken for comments on the manuscript, and they also thank A. Rodriguez-Navarro for discussing his research results prior to publication.

¹The National Technology Roadmap for Semiconductors (Semiconductor Industry Association, San Jose, CA, 1994).

²J. Ryan, R. M. Geffken, N. R. Poulin, and J. R. Paraszczak, IBM J. Res. Dev. **39**, 371 (1995).

³D. B. Knorr, Mater. Res. Soc. Symp. Proc. **309**, 75 (1993).

⁴D. P. Tracy, D. B. Knorr, and K. P. Rodbell, J. Appl. Phys. **76**, 2671 (1994).

⁵S. Vaidya and A. K. Sinha, Thin Solid Films **75**, 253 (1981).

⁶D. B. Knorr, D. P. Tracy, and K. P. Rodbell, Appl. Phys. Lett. **59**, 3241 (1991).

⁷D. B. Knorr and K. P. Rodbell, J. Appl. Phys. **79**, 2409 (1996).

⁸J. L. Hurd, K. P. Rodbell, D. B. Knorr, and N. L. Koligman, Mater. Res. Soc. Symp. Proc. **343**, 653 (1994); W. C. Shih, A. Ghiti, K. S. Low, A. L. Greer, A. G. O'Neill, and J. F. Walker, *ibid.* **428**, 249 (1996).

⁹J. M. E. Harper, E. G. Colgan, C. K. Hu, J. P. Hummel, L. P. Buchwalter, and C. E. Uzoh, Mater. Res. Bull. **19**, 23 (1994).

¹⁰D. P. Tracy and D. B. Knorr, J. Electron. Mater. **22**, 611 (1993).

¹¹E. M. Zielinski, R. P. Vinci, and J. C. Bravman, Mater. Res. Soc. Symp. Proc. **391**, 103 (1995).

¹²T. Nitta, T. Ohmi, M. Otsuki, T. Takewaki, and T. Shibata, J. Electrochem. Soc. **139**, 922 (1992).

¹³J. M. E. Harper, J. Gupta, D. A. Smith, J. W. Chang, K. L. Holloway, C. Cabral, Jr., D. P. Tracy and D. B. Knorr, Appl. Phys. Lett. **65**, 177 (1994).

¹⁴S. Q. Wang, Mater. Res. Bull. **19**, 30 (1994).

¹⁵H. J. Frost, Y. Hayashi, C. V. Thompson, and D. T. Walton, Mater. Res. Soc. Symp. Proc. **338**, 295 (1994).

¹⁶D. A. Porter and K. E. Easterling, *Phase Transformations in Metals and Alloys* (Van Nostrand Reinhold, UK, 1984).

¹⁷C. V. Thompson and R. Carel, Mater. Sci. Eng. B **32**, 211 (1995); J. A. Floro, R. Carel, and C. V. Thompson, Mater. Res. Soc. Symp. Proc. **317**, 419 (1994).

¹⁸*Recrystallization of Metallic Materials*, edited by F. Haessner (Dr. Riederer Verlag, Stuttgart, 1978).

¹⁹D. A. Smith and A. Ibrahim, Mater. Res. Soc. Symp. Proc. **317**, 401 (1994).

²⁰C. R. M. Grovenor, H. T. G. Hentzell, and D. A. Smith, Acta. Metall. **32**, 773 (1984).

²¹E. S. Machlin, *Materials Science in Microelectronics—The Relationships Between Thin Film Processing and Structure* (Giro, New York, 1995).

²²J. E. Sanchez, Jr., Mater. Res. Soc. Symp. Proc. **343**, 641 (1994).

²³J.-S. Yu, A. M. Maniatty, and D. B. Knorr, J. Mech. Phys. Solids (to be published).

²⁴E. M. Zielinski, R. P. Vinci, and J. C. Bravman, Mater. Res. Soc. Symp. Proc. **391**, 103 (1995).

²⁵C. V. Thompson, Mater. Res. Soc. Symp. Proc. **343**, 3 (1994).

²⁶G. A. Dixit, M. F. Chisholm, M. K. Jain, T. Weaver, L. M. Ting, S. Poarch, K. Mizobuchi, R. H. Havemann, C. D. Dobson, A. I. Jeffreys, P. J. Holverson, P. Rich, D. C. Butler, and J. Hems, Proceedings of the IEEE International Electron Devices Meeting, 1994, Vol. 94-105, p. 531.

²⁷G. Bai, C. Chiang, J. N. Cox, S. Fang, D. S. Gardner, A. Mack, T. Marieb, X.-C. Mu, V. Ochoa, R. Villasol, and J. Yu, 1996 IEEE Symposium on VLSI Technology Digest, 1996, p. 48.

²⁸M. J. C. van den Homberg, P. F. A. Alkemade, J. L. Hurd, G. J. Leusink, and S. Radelaar, Mater. Res. Soc. Symp. Proc. **391**, 397 (1995).

²⁹O. Kubachewski and C. B. Alcock, *Metallurgical Thermochemistry*, International Series on Materials Science and Technology Vol. 24, 5th ed. (Pergamon, Oxford, 1979).

³⁰A. Mineo, A. Matsuda, T. Kuroso, and M. Kikuchi, Solid State Commun. **13**, 329 (1973).

³¹Q. Z. Hong, F. M. d'Heurle, J. M. E. Harper, and S. Q. Hong, Appl. Phys. Lett. **62**, 2637 (1993).

³²C. Michaelsen, S. Wohler, and R. Borman, Mater. Res. Soc. Symp. Proc. **343**, 205 (1994).

³³L. A. Clevenger, C. V. Thompson, and K. N. Tu, J. Appl. Phys. **67**, 2894 (1990).

³⁴N. F. Mott and H. Jones, *The Theory of the Properties of Metals and Alloys* (Dover, New York, 1958), p. 23.

³⁵C. Cabra, Jr., K. Barmak, J. Gupta, L. A. Clevenger, B. Arcot, D. A. Smith, and J. M. E. Harper, J. Vac. Sci. Technol. B **11**, 1435 (1993).

³⁶L. A. Clevenger, A. G. Mutscheller, J. M. E. Harper, C. Cabral, Jr., and K. Barmak, J. Appl. Phys. **72**, 4918 (1992).

³⁷R. Beyers and R. Sinclair, J. Appl. Phys. **57**, 5240 (1985); F. M. d'Heurle, P. Gas, I. Engstrom, S. Nygren, M. Ostling, and C. S. Petersson, IBM RC Report No. 11151, Yorktown Heights, NY, 1985.

³⁸R. W. Mann, L. A. Clevenger, P. D. Agnello, and F. R. White, IBM J. Res. Dev. **39**, 403 (1995).

³⁹R. Jeanjean, J. Dubois, Y. Fétiveau, and R. Riviere, Mem. Sci. Rev. Met. **69**, 165 (1972).

⁴⁰I. J. M. M. Raaijmakers and K. B. Kim, J. Appl. Phys. **67**, 6255 (1990); I. J. M. M. Raaijmakers, Ph.D. thesis, Technical University of Eindhoven, 1988, p. 76 (unpublished).

⁴¹J. M. E. Harper, J. J. Cuomo, R. J. Gambino, and H. R. Kaufman, in *Surface Modification by Ion Bombardment: Fundamentals and Applications*, edited by R. J. Kelly and O. Auciello (Elsevier, Amsterdam, 1984).

⁴²E. Kay and S. M. Rossnagel, in *Handbook of Ion Beam Processing Technology*, edited by J. J. Cuomo, S. M. Rossnagel, and H. R. Kaufman (Noyes, Park Ridge, NJ, 1989), p. 170.

⁴³F. A. Smidt, Int. Mater. Rev. **35**, 61 (1990); P. J. Martin, J. Mater. Sci. **21**, 1 (1986).

⁴⁴J. E. Greene, in *Handbook of Deposition Technologies for Films and Coatings*, 2nd ed., edited by R. F. Bunshah (Noyes, Park Ridge, NJ, 1994).

⁴⁵L. S. Yu, J. M. E. Harper, J. J. Cuomo, and D. A. Smith, J. Vac. Sci. Technol. A **4**, 443 (1986); L. S. Yu, J. M. E. Harper, J. J. Cuomo, and D. A. Smith, Appl. Phys. Lett. **47**, 932 (1985).

⁴⁶H. Ji, G. S. Was, and J. W. Jones, Mater. Res. Soc. Symp. Proc. (to be published); J. J. Cuomo, C. R. Guarneri, R. H. Hammond, J. M. E. Harper, S. Herd, and D. S. Yu, IBM Tech. Discl. Bull. **25**, 3331 (1982).

⁴⁷R. E. Reed-Hill, *Physical Metallurgy Principles* (Van Nostrand, New York, 1973).

⁴⁸R. Venkatraman, J. C. Bravman, W. D. Nix, P. W. Davies, P. A. Flinn, and D. B. Fraser, J. Electron. Mater. **19**, 1231 (1990).

⁴⁹A. Goyal, D. P. Norton, J. D. Budai, M. Paranthaman, E. D. Specht, D. M. Kroeger, D. K. Christen, Q. He, B. Saffian, F. A. List, D. F. Lee, P. M. Martin, C. E. Klabunde, E. Hartfield, and V. K. Sikka, Appl. Phys. Lett. **69**, 1795 (1996).

⁵⁰J. Koike, P. R. Okamoto, L. E. Rehn, R. Bhadra, M. H. Grimsditch, and M. Meshii, Mater. Res. Soc. Symp. Proc. **157**, 777 (1990).

⁵¹G. L. Olson and J. A. Roth, Mater. Sci. Rep. **3**, 1 (1988).

- ⁵²H. A. Atwater, C. V. Thompson, and H. I. Smith, *J. Appl. Phys.* **64**, 2337 (1988).
- ⁵³H. Onoda, K. Touchi, and K. Hashimoto, *Jpn. J. Appl. Phys.* **1** **34**, 1037 (1995).
- ⁵⁴K. P. Rodbell, V. Sivilan, L. M. Gignac, P. W. DeHaven, R. J. Murphy, and T. J. Licata, *Mater. Res. Soc. Symp. Proc.* **428**, 261 (1996).
- ⁵⁵R. Murphy and K. P. Rodbell (unpublished).
- ⁵⁶P. Bai, J. F. McDonald, T.-M. Lu, and M. J. Costa, *J. Vac. Sci. Technol. A* **9**, 2113 (1991).
- ⁵⁷K. P. Rodbell, J. L. Hurd, and P. W. DeHaven, *Mater. Res. Soc. Symp. Proc.* **403**, 617 (1996).
- ⁵⁸P. R. Besser, J. E. Sanchez, Jr., and D. Field, *Proceedings of the Conference on Advanced Metallization and Interconnect Systems for ULSI Applications* (to be published).
- ⁵⁹J. M. E. Harper, in *Plasma-Surface Interaction and Processing of Materials*, edited by O. Auciello (Kluwer Academic, Dordrecht, The Netherlands, 1990), p. 251.
- ⁶⁰D. Dobrev, *Thin Solid Films* **92**, 41 (1982).
- ⁶¹D. D. Wang and T. Oki, *J. Vac. Sci. Technol. A* **8**, 3163 (1990).
- ⁶²I. Petrov, F. Adibi, J. E. Greene, L. Hultman, and J.-E. Sundgren, *J. Appl. Phys.* **73**, 8580 (1993).
- ⁶³Y. W. Kim, I. Petrov, J. E. Greene, and S. M. Rossnagel, *J. Vac. Sci. Technol. A* **14**, 346 (1996).
- ⁶⁴J. M. E. Harper, D. A. Smith, L. S. Yu, and J. J. Cuomo, *Mater. Res. Soc. Symp. Proc.* **51**, 343 (1986).
- ⁶⁵R. M. Bradley, J. M. E. Harper, and D. A. Smith, *J. Appl. Phys.* **60**, 4160 (1986).
- ⁶⁶Y. Iijima, K. Onabe, N. Futaki, N. Sadakata, O. Kohno, and Y. Ikeno, *J. Appl. Phys.* **74**, 1905 (1993).
- ⁶⁷N. Sonnenberg, A. S. Longo, M. J. Cima, B. P. Chang, K. G. Ressler, P. C. McIntyre, and Y. P. Lu, *J. Appl. Phys.* **74**, 1027 (1993).
- ⁶⁸P. McIntyre, K. G. Ressler, N. Sonnenberg, and M. J. Cima, *J. Vac. Sci. Technol. A* **14**, 210 (1996).
- ⁶⁹D. W. Hoffman and J. A. Thornton, *J. Vac. Sci. Technol.* **16**, 134 (1979).
- ⁷⁰J. M. E. Harper, K. P. Rodbell, E. G. Colgan, and R. H. Hammond, *Mater. Res. Soc. Symp. Proc.* (to be published).
- ⁷¹A. K. Malhotra, S. M. Yalisove, and J. C. Bilello, *Mater. Res. Soc. Symp. Proc.* **403**, 33 (1996); O. P. Karpenko, J. C. Bilello, and S. M. Yalisove, *J. Appl. Phys.* **76**, 4610 (1994); R. W. Smith, F. Ying, and D. J. Srolovitz, *Mater. Res. Soc. Symp. Proc.* **403**, 39 (1996).
- ⁷²S. Kominami, R. Mita, and T. Nishino, *J. Vac. Sci. Technol. B* **14**, 3395 (1996).
- ⁷³A. Rodriguez-Navarro, W. Otano-Rivera, J. M. Garcia-Ruiz, R. Messier, and L. J. Pilione, *J. Mater. Res.* (to be published).
- ⁷⁴H. Kreye and E. Hornbogen, *J. Mater. Sci.* **5**, 89 (1970).
- ⁷⁵M. F. Ashby, J. M. E. Harper, and J. Lewis, *Trans. Metall. Soc. AIME* **245**, 413 (1969).
- ⁷⁶C. V. Thompson, *J. Appl. Phys.* **58**, 763 (1985).
- ⁷⁷H. J. Frost, Y. Hayashi, C. V. Thompson, and D. T. Walton, *Mater. Res. Soc. Symp. Proc.* **317**, 431 (1994).
- ⁷⁸L.-C. Yang and A. Rockett, *J. Appl. Phys.* **75**, 1185 (1994).
- ⁷⁹L. Gignac, V. Sivilan, L. A. Clevenger, C. Cabral, Jr., and C. Lavoie, *Mater. Res. Soc. Symp. Proc.* (to be published).
- ⁸⁰R. A. Roy, L. A. Clevenger, C. Cabral, Jr., K. L. Saenger, J. L. Jordan-Sweet, J. Bucchignano, G. Morales, K. F. Ludwig, Jr., and G. B. Stephenson, *Appl. Phys. Lett.* **66**, 1732 (1995).
- ⁸¹V. Sivilan, K. P. Rodbell, L. A. Clevenger, C. Cabral, Jr., R. A. Roy, C. Lavoie, J. L. Jordan-Sweet, and J. M. E. Harper, *Mater. Res. Soc. Symp. Proc.* **427**, 53 (1996).
- ⁸²*Diffusion Phenomena in Thin Films and Microelectronic Materials*, edited by D. Gupta and P. S. Ho (Noyes, Park Ridge, NJ, 1988), p. 43.
- ⁸³N. A. Gjostein, *Diffusion* (American Society for Metals, Metals Park, OH, 1973), p. 241.
- ⁸⁴Z. Ma, L. A. Allen, and D. D. J. Allman, *J. Appl. Phys.* **77**, 4384 (1995).
- ⁸⁵J. Lasky, J. Nakos, O. Cain, and P. Geiss, *IEEE Trans. Electron Devices* **ED-38**, 2629 (1991).

Dual unit scanning tunneling microscope-atomic force microscope for length measurement based on reference scales

Haijun Zhang^{a)} and Feng Huang

Department of Optical Instruments, Zhejiang University, Hangzhou 310027, People's Republic of China

Toshiro Higuchi^{b)}

Kanagawa Academy of Science and Technology, 3-2-1 Sakado, Takatsu-ku, Kawasaki 213, Japan

(Received 8 January 1997; accepted 4 April 1997)

The crystalline lattice and periodic grating features have been employed as reference scales for simultaneous calibration between the scanning tunneling microscope (STM) and atomic force microscope (AFM). This article describes the configuration and functions of a dual unit STM-AFM for length metrology. The system was developed on the basis of a dual tunneling unit STM (DTU-STM) in view of the fact that the application of DTU-STM was strongly limited by sample conductivity. In the dual unit STM-AFM, a STM serving as a reference unit and an AFM as a test unit were combined by using one single *XY* scanner. Both the reference sample and the test sample were installed at the center of the scanner on either surface, and were imaged by STM and AFM units at the same time. The length of the test sample image was measured by counting the periodic features of the reference sample image. We present a detailed discussion about the structure and control of the dual unit STM-AFM. Some comparison results, respectively, using crystalline lattices and periodic grating features as reference scales for length calibration of test samples are also provided. Experiments show a satisfactory matching between the STM unit and the AFM unit when covering a wide scan range from 5 nm to 10 μ m. Using different standard reference scales, the system enables nanometer and/or sub-micron accuracy length metrology of microstructures with any conductivity. © 1997 American Vacuum Society. [S0734-211X(97)00304-1]

I. INTRODUCTION

For precise length metrology using scanning probe microscopes,^{1,2} a well-defined reference scale is needed to calibrate the scanning movement of the piezoscanner. The crystalline lattice of highly oriented pyrolytic graphite (HOPG) has been successfully employed as such a scale for simultaneous calibration in scanning tunneling microscopes (STMs),^{3,4} and some dual tunneling-unit STMs (DTU-STMs) have been developed for lateral^{5,6} and/or vertical⁷ scale calibration of the samples. In a DTU-STM, the test sample and HOPG chip are scanned by using one single *XY* scanner so that their images have the same size. The length in the test sample image is measured by counting the number of reference HOPG lattices. To automatically measure the length of the reference image, software developed for counting the number of lattices has also been presented.⁸

Recently, we have been using crystalline lattices of HOPG and periodic features of grating as reference scales in STMs, and constructed a new version of DTU-STM for practical application.⁹ Despite their excellence, however, an obvious limitation exists in DTU-STMs, that is, they are only capable of measuring electrically conductive samples. Before surface measurements of non-conductor samples with DTU-STMs, a thin film of conductive material should be deposited

onto the sample surface. This process is troublesome and undoubtedly changes the original structure of the sample surface. For direct measurement of non-conductor samples, we have been considering the replacement of the test tunneling unit of DTU-STM by an atomic force microscope (AFM) unit, and developed a dual unit STM-AFM for length measurement based on standard reference scales. In this article, we describe the configuration and functions of the newly developed STM-AFM system. Some comparison measurement results for length calibration of test samples using crystalline lattices and grating features as reference scales are also presented.

II. THE INSTRUMENT

A. Measurement concept

Figure 1 shows the measurement concept of the dual unit STM-AFM based on the design of a DTU-STM. One STM unit is kept as a reference unit but the other one is replaced by an AFM unit to image test samples. Samples with periodic features such as crystalline lattices and gratings are used as reference scales, respectively, for nanometer and micron order length measurement. The STM image of the reference sample and AFM image of the test sample are simultaneously scanned by one single *XY* scanner. They are thus of the same lateral size, and the length in the test sample image could be measured by counting the number of periodic features in the reference image.

^{a)}Also at Kanagawa Academy of Science and Technology, 3-2-1 Sakado, Takatsu-ku, Kawasaki 213, Japan.

^{b)}Also at Department of Precision Machinery Engineering, the University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan.

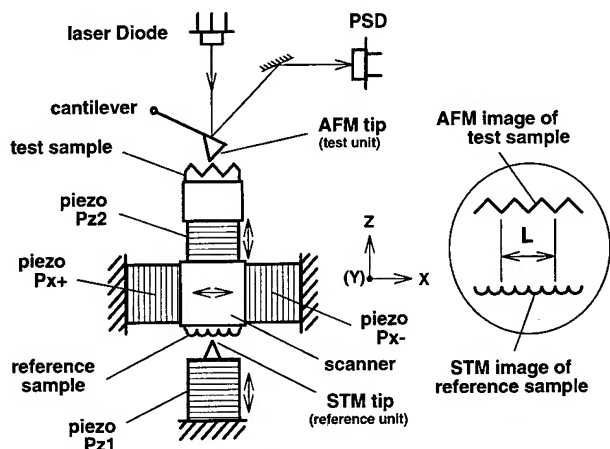


FIG. 1. Measurement concept of length metrology based on reference scale by a dual unit STM-AFM. The length of test image is measured by counting the number of periodic features of reference image.

B. Structure of the dual unit STM-AFM

The schematic diagram of the dual unit STM-AFM is shown in Fig. 2. The STM unit and AFM unit were vertically set in series. The reference sample holder and test sample holder with piezo P_{z2} were attached to the center of the scanner on either surface. The optical beam deflection (OBD) method^{10,11} was applied to detect the deflection of the cantilever caused by atomic force between the AFM tip and test sample. A laser beam was focused on the cantilever and then reflected onto the surface of a position sensitive detector (PSD) at a distance of 10 cm from the cantilever. The output difference of photoelectric current from PSD was related to the deflection of the cantilever. The length of the cantilever is 200 or 100 μm . As a result, the deflection signal could be amplified by 1000 or 2000 times. A charge-coupled device

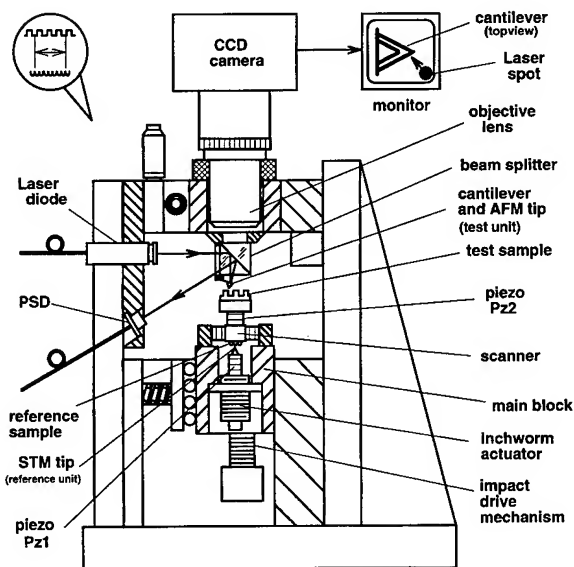


FIG. 2. Schematic diagram of the dual unit STM-AFM, incorporating a STM reference unit and an AFM test unit sharing one single XY scanner.

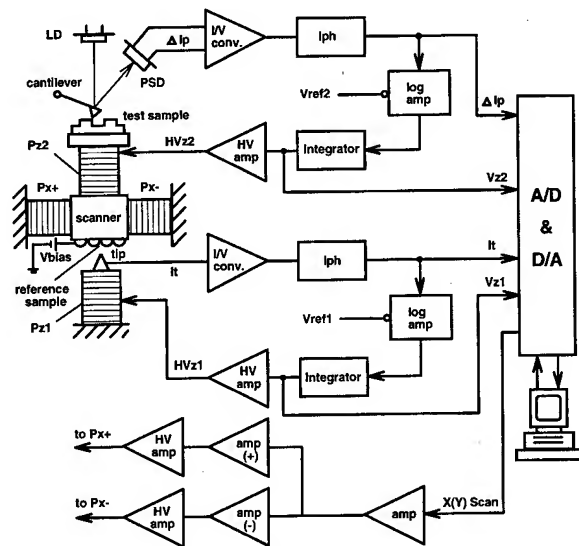


FIG. 3. Block diagram of the PID feedback control circuits for the two units of the dual unit STM-AFM, and scan control circuit of the XY scanner for the X direction, same for the Y direction.

(CCD) camera with an optical microscope was used for convenient and quick positioning of a focused laser spot onto the cantilever.

To bring the reference unit into the tunneling controlled state, an inchworm actuator was used to coarsely approach the STM probe tip. The actuator was stopped when tunneling current increased to a given value. The AFM unit was brought into the atomic force controlled state by using an impact drive mechanism (IDM).¹² The test sample was pushed to move vertically together with the whole main block, close to the fixed cantilever. The IDM was stopped when the photoelectric current difference from PSD increased to a given value. The movement sequence of the inchworm actuator and IDM has been discussed in Ref. 9.

C. Feedback control of the Z piezos

The probe tip of the STM unit and the test sample of the AFM unit were precisely positioned in the Z direction by using piezos P_{z1} and P_{z2} . Figure 3 shows the block diagram of two feedback loops for controlling the movement of these piezos. The main part of each loop is a conventional proportional-integral-differential (PID) feedback circuit. Take the STM loop, it consists of a tunneling current to voltage convertor, a low-pass filter, a logarithmic amplifier, and an integrator. The voltage signal corresponding to tunneling current is compared with a reference voltage V_{ref1} . Their difference is then fed to the integrator which provides a positive or negative voltage difference to a high-voltage amplifier. The voltage difference controls the movement of P_{z1} to change the tunneling gap as to increase or decrease the tunneling current I_t . In this way, the stable tunneling state is automatically maintained.

The AFM loop is similar to the STM loop, except that it picks the photoelectric current difference from PSD pins instead of tunneling current from the STM tip. Simulation was

carried out to determine favorable conditions of the feedback parameters for the two units. Noise tolerance was reduced to the orders in which, for the STM unit, high contrast crystalline lattices could be repeatably observed in raw atomic images, and for the AFM unit, nanometer order structures in raw images.

Both of the PID loops provide two feedback modes: tunneling current or atomic force (correspondingly, photoelectric current) constant mode, and tip or sample height constant mode. We used the first mode for precisely positioning tips and samples in the Z direction. When the two units were brought into stable tunneling current and/or atomic force states, we turned to employ the second mode for image acquisition. The main considerations are: (i) if the axes of P_{Z1} and P_{Z2} are not parallel to each other; and (ii) if they are not perpendicular to the XY plane, the movement of P_{Z1} and P_{Z2} might cause additional errors to XY scales of the images. Therefore, we kept the heights of P_{Z1} and P_{Z2} to be constant during image acquisition. The XY scales of reference and test images were then uniquely determined by the scanner and were exactly the same.

Four channels of A/D convertors were used to acquire image data: tunneling current I_t and photoelectric current difference ΔI_p , and additionally piezo voltages V_{Z1} and V_{Z2} .

D. Control of the XY scanner

Within an XY scanner, if there is an interference between the X and Y movement of the piezos, the lateral scales of images might be distorted. In the dual unit STM-AFM, we used an elastic XY scanner to avoid lateral image distortion (Ref. 9). The block diagram of scan control circuit for the X direction is also shown in Fig. 3, the same for the Y direction. Scanning signals were generated by a personal computer and fed to the circuit through a D/A interface. The circuit enables image acquisition covering a scan range from 5 nm to 10 μm .

III. EXPERIMENTS

Experiments were carried out on the dual unit STM-AFM to check the feasibility of simultaneous calibration between STM and AFM. A HOPG chip was generally used as a reference sample for precise length metrology. Before each experiment, the HOPG surface was freshly cleaved. STM tips were prepared from mechanically cut 0.25 mm platinum/iridium wire. AFM cantilevers were the commercial products of Seiko Instruments Corporation. A 2.5 mW laser diode was used in the AFM unit, wavelength $\lambda=670$ nm. The typical scanning conditions were, for the STM unit, $I_{\text{ref}}=-1.0$ nA, $V_{\text{bias}}=+50$ mV (sample positive), tip height constant mode, and for the AFM unit, $V_{\text{ref2}}=+0.5$ V, sample height constant mode. All images were taken in air and under a temperature condition of $23^\circ\text{C}\pm 0.1^\circ\text{C}$.

Considerations for the software of image acquisition are as follows: (i) the reference sample and test sample are simultaneously scanned and imaged; (ii) because the two samples are installed back to back on the scanner, and two

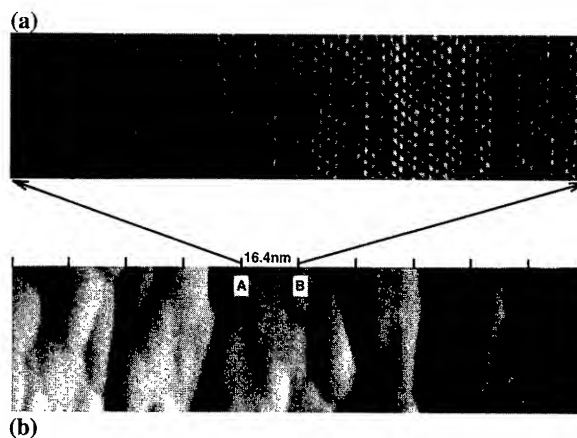


FIG. 4. An example of comparison measurement by using the dual unit STM-AFM. Ten equal parts of STM images of HOPG lattices were acquired within the length of one AFM image of a gold-plated glass surface. The length of the fifth image part was counted to be 16.4 nm, equivalent to line AB in test image. The added length of the whole test image was about 164 nm. The width of images could be calculated in the same way. All images were scanned in the Y direction, tip, and sample height constant mode. (a) STM image of the fifth part of HOPG lattices (reference side); (b) AFM image of a gold-plated glass surface (test side).

probe tips have a head to head setup, the lateral coordinates of reference image and test image are actually corresponding to each other in a mirror-image way (see Fig. 1). For a convenient comparison, we saved in advance the reference image in its mirror-image form to obtain a positive correspondence between reference image and test image; and (iii) for low accuracy measurement, each image is scanned and saved in one complete form, while for high accuracy, the reference image is divided into several corresponding and equal parts for convenient scanning, saving, and for possible lattices counting.

Figure 4 shows an example of comparison measurement using HOPG as a reference sample. The test target is a gold-plated glass surface. Ten equal parts of the STM images of crystalline lattices were obtained within the length of one AFM image of the test sample. Each image part contained 600×150 scan pixels, scanned in the Y direction. Their lengths were measured by counting the number of lattices. The length of the fifth reference part is equivalent to that of line AB in the test image, and is measured to be 16.4 nm. The other nine parts were omitted for clarity. The total length of the test sample image is added to be about 164 nm. The width of the test image can be calculated in the same way. Notice that the reference graphite lattice has a slight distortion which we need to eliminate. However, this sort of distortion will not directly affect the accuracy of distance measurements. This is because the image length is measured by counting the number of lattices. Even if the lattice image is slightly distorted, we still precisely know the distance between two adjacent lattices, and the added distance should also be accurate.

For wide range measurement up to microns, it is definite that we can still use the crystalline lattice as a reference scale, however, the image size of reference lattices will be

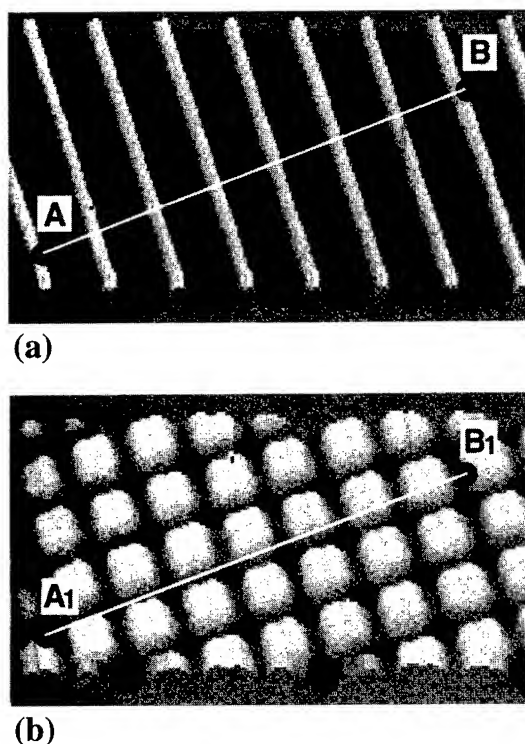


FIG. 5. Wide range calibration of the dual unit STM-AFM by using periodic grating (2000 lines/mm) and grating pattern (2000×2000 lines/mm²) as reference and test samples. The 360×180 images were scanned in the X direction, tip, and sample height constant mode, 1 line/s, 4.0 μm ×2.0 μm scan. The lengths of corresponding lines AB and A₁B₁ were 3.5 and 3.6 μm , with a difference of 2.8%. The results show that the STM and AFM images have a satisfactory matching of lateral scales. (a) STM image of grating (reference side); (b) AFM image of grating pattern (test side).

explosively large. In this case, we consider employing some other samples with periodic features such as standard gratings as reference scales. Figure 5 is an experimental result of wide range calibration between STM and AFM units. On the STM side, a gold-plated grating of 2000 lines/mm served as a reference sample. On the AFM side, a non-conductive grating pattern of 2000×2000 lines/mm² was used as a test sample. The two samples were aligned in advance to approximately obtain the same grating orientation. The images contained 360×180 scan pixels and were scanned under voltages 180 V×90 V with a slow frequency of 1 line/s in the X direction. The grey scales of the two images were already smoothed without changing their lateral scales. Again, the scan area of both images could be measured by counting the number of grating periods, and were approximately 4.0 μm ×2.0 μm . Take corresponding lines AB and A₁B₁ in the two images, their measured lengths were 3.5 and 3.6 μm , with a length difference of 2.8%. It is clear that the lateral scales of the two images match each other quite well. Length differences larger than 10% have even been observed when the sample surface of the grating pattern on the AFM side was not as parallel to that of grating in the STM side, but these errors were not caused by the system itself. Scan ranges up to 10 μm were realized by increasing scan voltage to 400 V.

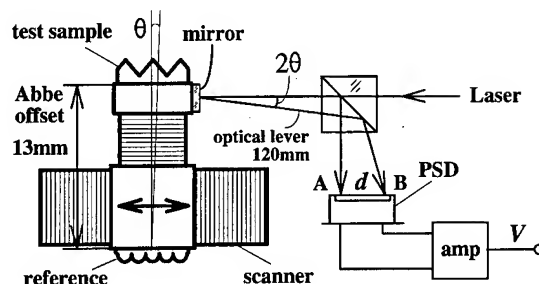


FIG. 6. Tests of the tilt angles of the central block by using the optical beam deflection method. The measured errors between STM and AFM images caused by Abbe offset were less than 2.47% and 1.62% in the X and Y directions, 10 μm ×10 μm scan.

In the apparent setup, the Abbe offset of 13 mm between the reference and unknown samples has a potential of causing errors to STM and AFM images. However the actual errors observed in the experiments were small. We confirmed that the central block was not twisting by two facts. First, the scanned grating features were quite regular as shown in Fig. 5, otherwise their images would be seriously distorted. Second, some tests were performed to determine the tilt angles of the central block by using the OBD method (Fig. 6). A mirror was mounted on the test sample stage. A focused laser beam incident on the mirror in the horizontal direction was perpendicularly reflected onto the PSD surface. This setup avoids the scanning movement of the scanner and is only sensitive to the tilt angle θ . For the wide scan range of 10 μm ×10 μm , the maximum tilt angles measured in the X and Y directions during image scanning were 0.019 and 0.0125 mrad. Due to the Abbe offset, these angles could cause image errors of 0.247 and 0.162 μm between the reference and unknown samples. The percentages of errors between them were 2.47% and 1.62% in the X and Y directions. These errors were probably the main cause of the relatively large 2.8% error detected when scanning identical gratings. For scan range from 5 nm to 1 μm , no twisting of the central block was detected. Even if the errors were tolerable, as a future study, we are planning to develop a new setup to avoid the potential influence of the Abbe offset.

Other causes of image errors may include the coupling of the X and Y piezos of the scanner and the coupling of the movement of Z piezos with the X and Y axes. By using the elastic XY scanner and taking sample and tip height constant modes during image acquisition, these errors can be successfully avoided and were practically eliminated.

IV. CONCLUSIONS

The dual unit STM-AFM system is an important addition to DTU-STM for length metrology of microstructures based on standard reference scales. It keeps the remarkable advantages of both STM and AFM, that is, the AFM unit is able to measure test samples with any electric conductivity, while the STM unit provides high resolution for reference images. Experiments show that the STM-AFM enables nanometer accuracy length measurement when employing the HOPG

crystalline lattice as a reference scale. For wide range measurement, satisfactory calibration between STM and AFM was realized by using the periodic grating features as reference scales. Scan ranges from 5 nm up to 10 μm were confirmed by experiments. This provides the possibility of practical observation and measurement for microstructures with different length scales. One disadvantage of the dual unit STM-AFM is that the resolution of the AFM unit is not as high as that of the STM unit. However, the system still maintains sufficient accuracy for length metrology from nanometer to micron orders.

ACKNOWLEDGMENTS

H. Zhang and F. Huang wish to thank the financial support from the Chinese National Foundation of Natural Sciences (Grant No. 69606003), and from the Foundation of Natural Sciences of Zhejiang Province, China (Grant No.

596095). This work was also supported by Kanagawa Academy of Science and Technology, Japan.

¹G. Binnig and H. Rohrer, *Helv. Phys. Acta* **55**, 726 (1982).

²G. Binnig and C. F. Quate, *Phys. Rev. Lett.* **56**, 930 (1986).

³E. C. Teague, *J. Vac. Sci. Technol. B* **7**, 1898 (1989).

⁴H. Kawakatsu and T. Higuchi, *J. Vac. Sci. Technol. A* **8**, 319 (1990).

⁵H. Kawakatsu, Y. Hoshi, and T. Higuchi, *J. Vac. Sci. Technol. B* **9**, 651 (1991).

⁶H. Kawakatsu, T. Higuchi, H. Kougami, M. Kawai, M. Watanabe, Y. Hoshi, and N. Nishioki, *J. Vac. Sci. Technol. B* **12**, 1681 (1994).

⁷T. Fujii, M. Suzuki, T. Higuchi, H. Kougami, and H. Kawakatsu, *J. Vac. Sci. Technol. B* **13**, 1112 (1995).

⁸H. Kawakatsu and H. Kougami, *J. Vac. Sci. Technol. B* **14**, 11 (1996).

⁹H. Zhang, T. Higuchi, and N. Nishioki, *J. Vac. Sci. Technol. B* **15**, 1 (1997).

¹⁰G. Meyer and N. M. Amer, *Appl. Phys. Lett.* **53**, 2588 (1988).

¹¹S. Alexander, L. Hellemans, O. Marti, J. Schneir, V. Elings, P. K. Hansma, M. Longmire, and J. Gurley, *J. Appl. Phys.* **65**, 164 (1988).

¹²T. Higuchi, in *Proceedings of the 2nd International Symposium of Robotics Research*, 1984 (unpublished), p. 144.

Magnetic nanostructures fabricated by scanning tunneling microscope-assisted chemical vapor deposition

Woei Wu Pai,^{a)} Jiandi Zhang, and John F. Wendelken
Solid State Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6057

R. J. Warmack
Health Sciences Research Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6123

(Received 28 January 1997; accepted 4 April 1997)

We have successfully used scanning tunneling microscope-assisted chemical vapor deposition to fabricate magnetic nanostructures as fine as 5 nm wide and <2 nm high using ferrocene [$\text{Fe}(\text{C}_5\text{H}_5)_2$] as the metal-organic source gas. The physical properties of these nanostructures were qualitatively characterized and *ex situ* magnetic force microscopy measurements indicate these features are strongly magnetic. [S0734-211X(97)00604-5]

I. INTRODUCTION

Scanning probe microscopes (SPMs) have been used to pattern features down to the nanoscale regime, below the limits obtainable by conventional methods such as photon and e-beam lithography. The techniques that have been used in SPM-assisted nanofabrication include field evaporation,¹ chemical vapor deposition (CVD),² resist patterning,^{3,4} etc. Among these fabrication techniques, scanning tunneling microscope (STM)-assisted CVD has been used to deposit Fe, Pd, Ni, Cd, Si, and Al.^{2,5-9} This article will describe the fabrication of Fe-containing magnetic nanostructures using STM-assisted CVD in ultrahigh vacuum (UHV) with [$\text{Fe}(\text{C}_5\text{H}_5)_2$] as the source gas at room temperature. Qualitative measurement of the magnetic properties of these features with an ambient magnetic force microscope (MFM) will also be presented. Generally, organometallic sources may be categorized into end-bound and side-bound ligand complexes.¹⁰ Previous studies of STM-induced fabrication of Fe-containing magnetic nanostructures have been made using end-bound ligand organometallic complexes such as iron-pentacarbonyl [$\text{Fe}(\text{CO})_5$].^{5,11} We use the side-bound organometallic complex ferrocene [$\text{Fe}(\text{C}_5\text{H}_5)_2$], as the source gas. Metallocenes, especially ferrocene, are known to possess remarkable thermal stability at room temperature but they become unstable at higher temperature, and thus are used in many CVD processes.^{12,13} The planar cyclopentadienyl (C_5H_5) ligand is also one of the most common ligands in organometallic chemistry.¹⁴ Demonstration of ferrocene as an effective source for SPM-based fabrication of magnetic nanostructures is thus of interest.

II. EXPERIMENT

The fabrication was performed on Ag(100) and Cu(100) single crystals in an UHV chamber having a base pressure $\sim 1 \times 10^{-10}$ Torr. A commercial Omicron STM was used with etched W tips. The Ag and Cu substrates were cleaned by repeated Ne sputtering and annealing cycles. During fab-

rication, a gas doser was positioned to within ~ 5 mm of the STM tip and surface gap regions. The pressure in the gas doser was measured independently using a baratron gauge and the pressure in the tip-surface region was estimated to be $\sim 1/5$ of that in the gas doser. Typically, the gas doser pressure used for fabrication ranges from 1 to 30 mTorr while the vacuum chamber pressure increased to $\leq 3 \times 10^{-7}$ Torr. The STM was operated in the field emission regime with a sample bias between 6 and 20 V and a feedback current between 0.2 and 20 nA. It was found that the sizes of the fabricated features are most sensitive to the magnitude of sample bias and are less sensitive to the feedback currents used. An ambient MFM was used to characterize *ex situ* the Fe-containing nanostructures. In order to protect the nanostructures against oxidation after removal from the UHV system, an approximately 50 monolayer (ML) film of the same material as the substrate was evaporated onto the nanostructures to form a protective cap layer before transferring them to the MFM. Nanostructures were fabricated in a region marked by two parallel lines having a separation of ~ 40 μm . After fabrication, the lines marked allowed relocation of the fabricated nanostructures in the MFM measurement.

III. RESULTS AND DISCUSSION

Figure 1 shows the best lines fabricated to date. The lines on Ag were produced by scanning the STM over the same line ~ 300 times in a 140 s period with a sample bias ranging from 7 to 10 V, a feedback current of 0.2 nA, and a gas line source pressure of 5 mTorr. Between the fabrication of each line, the tip position was offset by 15 nm and then the subsequent image was taken with a scan direction at 90 degrees relative to the fabricated lines. In Fig. 1(a), one thick line and one thin line were fabricated. In Figs. 1(b)–1(c), subsequent fine lines were made. Figure 1(d) shows the profile across the lines of Fig. 1(c). These lines can be as narrow as 5 nm [full width at half-maximum (FWHM)] and less than 2 nm in height with an interline spacing as close as 15 nm. The widths of these sub-10 nm lines approach the simulated beam size limit of field emission from the STM tip under similar conditions.¹⁵ We note that the thickness varies by as

^{a)} Author to whom correspondence should be addressed; Electronic mail: yie@ornl.gov

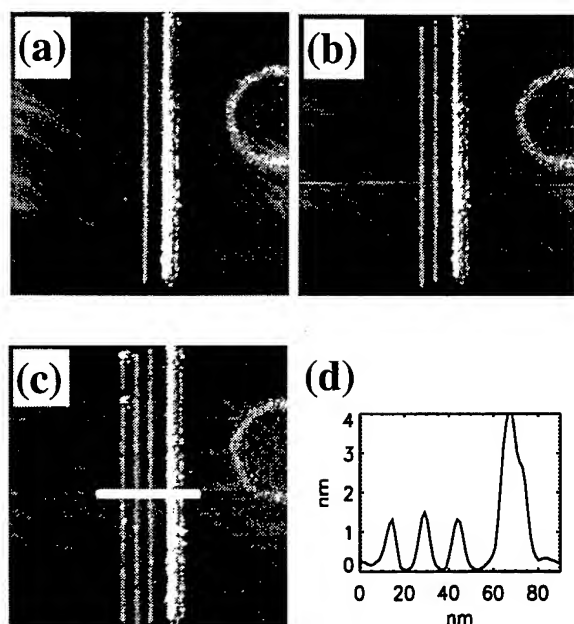


FIG. 1. Three consecutive STM images, (a)–(c), showing the capability for fabrication of closely spaced nanowires. The image size is $300\text{ nm} \times 300\text{ nm}$ and the substrate is Ag(100). Fine nanolines are fabricated with a sample bias of 7 V and, the thicker line with 10 V. Other parameters are 0.2 nA feedback current, 5 mTorr ferrocene gas line pressure, and 300 repetitive passes in 140 s. The upper right half-circle is also a fabricated feature. A line profile across these fine lines is shown in (d). The lines can be as narrow as 5 nm (FWHM) and less than 2 nm in height with an interline spacing as small as 15 nm.

much as 30% along the line. This is an undesirable and, unfortunately, general phenomenon observed in this as well as in other work using the STM-CVD approach. The roughness that occurs is reminiscent of the nucleation and growth of discrete islands which are often observed in the epitaxial growth of thin films. In this case, the role of diffusion in the growth process is supplemented, and perhaps overwhelmed, by an electric field effect between the tip and the surface. Once an island is nucleated in this process, a stronger field is created between the tip and the island site which then guides electrons or ionized atoms to this site to continue growth of the island.^{1,5} As the tip motion carries it further from the island, the field is minimized and the process begins again at a new site. Such an effect is evident in fabricating complex patterns in a single pass such as shown in Fig. 2. In fabricating this feature, the tip was programmed to follow a predefined path, and the fabrication voltage and current were kept fixed with the feedback loop on. Individual mounds were evident with a root mean square roughness of about $\sim 80\%$ of the average height. In general, one can minimize this roughness by using the lowest possible voltage to break the CVD source molecules at the initial stage of deposition.

The quality of the fabricated features depends on various writing conditions such as sample bias, feedback current, source pressure, and writing speed. We summarize the general fabrication characteristics as follows. (1) Deposition occurs at both polarities of sample bias, but a positive sample

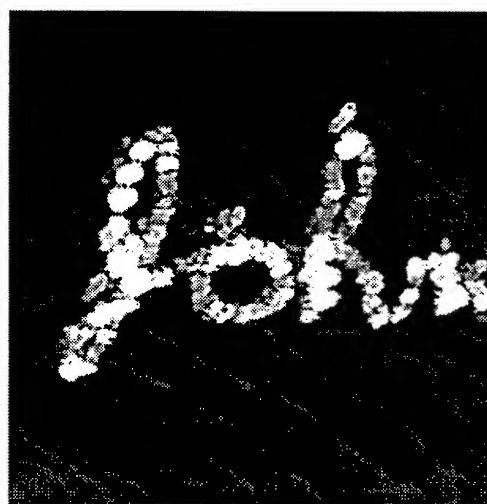


FIG. 2. By moving the tip along a predefined path, a complex arbitrary pattern can be fabricated. This $150\text{ nm} \times 150\text{ nm}$ image shows the signature of one of the authors fabricated in a single pass. Note the substrate (Ag) features such as steps remain intact after fabrication. The parameters are fabrication bias 6.6 V, feedback current 1.0 nA, gas line pressure 4.7 mTorr, and writing time 60 s.

bias is more efficient. By comparing the height of the nanostructures using the same fabrication conditions except bias polarity, we found the fabrication efficiency at positive sample bias is about five times greater than that at negative sample bias. The dominant cause of the molecular fragmentation is most likely electron bombardment from field emission of the tip. The higher yield at positive sample bias may be due to breaking of molecules by electrons with higher kinetic energies at regions closer to the surface rather than to the tip. (2) Deposition rates increase noticeably with increasing bias voltage and dosing pressure but are less sensitive to the feedback current.

The usefulness of these nanostructures which may be custom fabricated by this as well as by other scanning probe techniques is to provide systems for which novel physical properties can be explored and characterized. The extraordinarily small sizes of the fabricated features and the small numbers that can be realistically fabricated virtually rule out any characterization requiring the signals to be averaged over a region much larger than tens of nanometers. However, several groups have characterized these nanostructures either by using a scanning Auger microscope⁵ or use of probes such as a superconducting quantum interference device (SQUID) coil,¹⁶ or a Hall magnetometer.¹¹ In these measurements, the chemical compositions, hysteresis curve, or the collective magnetic excitations can be measured. One other method is to use magnetic force microscopy as the probe. Nevertheless, the rather poor resolution ($\geq 10\text{ nm}$) of a MFM compared to a STM imposes a limitation at this time. Furthermore, the usual ambient implementation of a MFM does not allow *in situ* measurement if the nanostructures were made in UHV. These limitations, however, can be overcome by careful preparation. We intentionally made larger features and, as

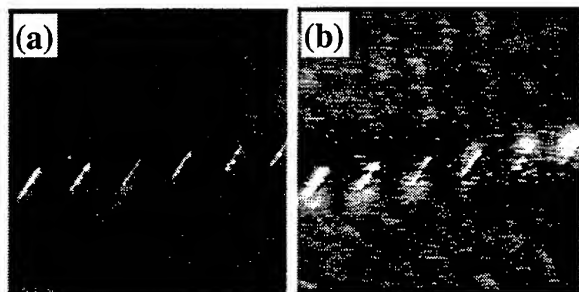


FIG. 3. Magnetic-force image ($8\ \mu\text{m} \times 8\ \mu\text{m}$) of an array of fabricated nanowires on Cu. These wires were fabricated in UHV, capped with Cu, and transferred to a MFM for *ex situ* measurement under ambient conditions. (a) The topography image showing five nanowires, $0.8\ \mu\text{m}$ long, $60\text{--}200\ \text{nm}$ wide and $50\text{--}80\ \text{nm}$ high. (b) Magnetic force image of the phase-shift signal (2° range) taken at a lift height of $80\ \text{nm}$. The interaction of the nanowires with the magnetic probe tip is apparent.

indicated above, carefully marked our nanofabrication region for positional identification and then capped the substrate and features with a $\sim 50\ \text{ML}$ protective layer. We have been able to successfully perform ambient MFM measurements, and the results are presented below.

Figure 3 shows such an *ex situ* ambient MFM measurement on an array of nanowires fabricated on Cu in UHV with a $20\ \text{V}$ sample bias and $0.2\ \text{nA}$ feedback current. These lines were made in larger sizes to facilitate finding them after repositioning within the MFM. Imaging was performed with a Digital Instruments Nanoscope III operated in the "lift mode" to separate the magnetic force signals from the topographic signals. In this mode, the tip first performs a tapping-mode, single-line topographic scan and then a second scan is made at a sufficiently large constant height relative to the measured topography to detect long-range interaction forces with the magnetic tip. Magnetic force was measured by the phase shift between the driving frequency and the modulation of collected signal intensity. The topographic map and magnetic force map are thus interleaved, line-by-line. Figure 3 shows a MFM image pair of a portion of the STM-fabricated array. The bright nanowires visible in the topographic image [Fig. 3(a)] are replicated in the phase image [Fig. 3(b)] taken with an $80\ \text{nm}$ lift. The contrast in this image corresponds to a phase shift of $1^\circ\text{--}2^\circ$ (comparable to that of magnetic tape recorded under similar conditions). A slowly decreasing contrast with lift height was seen to be $\sim 0.5^\circ$ at $230\ \text{nm}$, further confirming the presence of a

strong, long-range magnetic interaction. The fabricated lines appear to carry the signature of ferromagnetism of a polycrystalline material.

IV. CONCLUSION

We have successfully fabricated Fe-containing magnetic nanostructures using STM-assisted CVD with ferrocene as the source. The ambient magnetic force microscopy measurement indicates the structures are strongly magnetic. The ability to fabricate magnetic structures of controlled geometry down to the nanoscale regime will allow further systematic investigations of the physics of these nanoscale objects.

ACKNOWLEDGMENTS

The authors thank P. A. Dowben for providing the ferrocene source and for helpful discussions. This research was supported in part by an appointment to the Oak Ridge National Laboratory and the Oak Ridge Institute for Science and Education and by the U. S. Department of Energy, Division of Materials Sciences, under Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp.

¹C. S. Chang, W. B. Su, and T. T. Tsong, *Phys. Rev. Lett.* **72**, 574 (1994).

²R. M. Silver, E. E. Ehrichs, and A. L. de Lozanne, *Appl. Phys. Lett.* **51**, 247 (1987).

³C. Vanhaesendonck, L. Stockman, Y. Bruynseraede, L. Langer, V. Bayot, J. P. Issi, J. P. Heremans, and C. H. Olk, *Phys. Scr.* **55**, 86 (1994).

⁴D. P. Adams, T. M. Mayer, and B. S. Swartzentruber, *Appl. Phys. Lett.* **68**, 2210 (1996).

⁵A. D. Kent, T. M. Shaw, S. V. Molnar, and D. D. Awschalom, *Science* **262**, 1249 (1993).

⁶D. S. Saulys, A. Ermakov, E. L. Garfunkel, and P. A. Dowben, *J. Appl. Phys.* **76**, 7639 (1994).

⁷S. Rubel, X. D. Wang, and A. L. de Lozanne, *J. Vac. Sci. Technol. B* **13**, 1332 (1995).

⁸T. M. H. Wong, S. J. O'Shea, A. W. McKinnon, and M. E. Welland, *Appl. Phys. Lett.* **67**, 786 (1995).

⁹S. T. Yau, D. Saltz, A. Wriekat, and M. H. Nayfeh, *J. Appl. Phys.* **69**, 2970 (1991).

¹⁰J. T. Spencer, *Prog. Inorg. Chem.* **41**, 145 (1994).

¹¹A. D. Kent, S. V. Molnar, S. Gider, and D. D. Awschalom, *J. Appl. Phys.* **76**, 6656 (1994).

¹²F. Thibaudau, J. R. Roche, and F. Salvan, *Appl. Phys. Lett.* **64**, 523 (1994).

¹³D. Welipitiya, Y. L. He, Jiandi Zhang, P. I. Oden, T. Thundat, R. J. Warmack, I. Gobulokoglu, Z. S. Shan, D. J. Sellmyer, and P. A. Dowben, *J. Appl. Phys.* **80**, 1867 (1996).

¹⁴T. J. Marks, *Prog. Inorg. Chem.* **25**, 223 (1979).

¹⁵T. M. Mayer, D. P. Adams, and B. M. Marder, *J. Vac. Sci. Technol. B* **14**, 2438 (1996).

¹⁶D. D. Awschalom, M. A. McCord, and G. Grinstein, *Phys. Rev. Lett.* **65**, 783 (1990).

Force modulation atomic force microscopy recording for ultrahigh density recording

S. Hosaka,^{a)} H. Koyanagi, A. Kikukawa, M. Miyamoto, K. Nakamura, and K. Etoh
Advanced Research Laboratory, Hitachi Limited, 1-280 Higashi-koigakubo, Kokubunji, Tokyo 185, Japan

(Received 12 February 1997; accepted 16 May 1997)

We propose force modulation atomic force microscopy (FM)-(AFM) pit recording and demonstrate the possibility of achieving ultrahigh density recording with high speed readout. A minimum pit size of around 10 nm in diameter is formed by cold plastic deformation of the polycarbonate disk surface at a force of over 40 nN. Using a prototype of the rotation type FM-AFM pit recording system, an ultrahigh recording density of 1.2 Tb/in.² and a readout speed of 1.25 Mb/s are demonstrated in 1/2(2,7) code recording. © 1997 American Vacuum Society. [S0734-211X(97)13804-5]

I. INTRODUCTION

The areal recording density of data storage has increased 10- to 30-fold over the last 10 years. However, we predict that a physical limit will appear at several 10s Gb(giga bit)/in.² because of superparamagnetics in magnetic and optical diffraction in optical recordings. In order to overcome this limit, scanning probe microscope (SPM) based data storage is expected as one of future storages.

Many SPM based data storage schemes have been proposed with scanning tunneling microscope (STM), atomic force microscope (AFM), magnetic force microscope (MFM), and scanning near-field optical microscope (SNOM) techniques to achieve ultrahigh density recording with a bit size of less than 60 nm.¹⁻⁵ STM recording has the potential to achieve 6 Pb(peta bit)/in.² using an atom or a vacancy as an information unit. The AFM demonstrated a potential of achieving 1 Tb(trillion bit)/in.² while MFM or SNOM recordings with 100 Gb/in.² were demonstrated. However, these have been proposed only to achieve ultrahigh recording density. Their readout speeds are still far from the practical speed of faster than 1 Mb/s. In a previous article, we discussed the possibility of not only ultrahigh recording density but also high readout speed.³ For high speed readout, a gap control comparison using various SPM techniques discussed that AFM and SNOM based storages are very suitable because they have the possibility of achieving high frequency response of gap control or the redundancy to permit servo control error in the signal-gap properties.

Due to these considerations, we reached the conclusion that AFM based storage is a more adaptive method for densities higher than 1 Tb/in.² and readouts faster than 1 Mb/s. Recently, many AFM recordings using pit formation,⁶ dot formation,⁷ phase change,⁸ and charge control⁹ have been proposed by numerous researchers. The pit and the dot formation AFM recordings have higher reliability in writing and reading than phase change and charge control type AFM recordings. Considering the pit and dot formation AFM recordings, the former method is highly reliable for writing, but it has a problem with bit size in that the minimum bit size is larger than 100 nm in diameter. On the other hand, the

latter is capable of forming nanometer sized dots of less than 20 nm in diameter, but it has a problem in the stable formation of the dots. The nanoindentation technique using AFM was introduced early by Jung *et al.* and Hamada and Kaneko. They demonstrated nanopit formation with a dot of 50–100 nm in diameter.^{10,11} We applied this technique to ultrahigh density recording to achieve AFM pit recording using force modulation because it is 100% successful in pit formation with a nanometer size of less than 25 nm.

In this article, we describe force modulation AFM pit recording, and demonstrate that it has the potential to attain ultrahigh density recording with high speed readout.

II. FORCE MODULATION ATOMIC FORCE MICROSCOPY (FM-AFM) PIT RECORDING

The pits were made by cold plastic deformation (nano-indentation) for the writing process. This was done by pushing the AFM tip into a sample surface with the piezo actuator shown in Fig. 1(a). Using this method, the force was precisely controlled at a resolution of a few nN. Here, the sample was a polycarbonate disk.

For the reading process, we read pit structures with a laser beam deflection system the same as for AFM [Fig. 1(b)].^{12,13} When the tip runs on the sample surface at a high speed, the pit structures on the sample make the cantilever vibrate. This is because gap control is not responsive to high speed scanning (readout). In order to do such a experiment, we had to develop a high speed scanning method and high-speed-responses detection. This was mainly achieved by developing a rotation type AFM recording system and a high resonant frequency AFM cantilever.

III. NANOPIT FORMATION USING NANOFORCE MODULATION

Figure 2 shows a simple model and an equivalent model of Fig. 1(a). When the cantilever is pushed so that it is at a distance of X_0 from the sample, the tip penetrates the surface to a depth of X . The working force of the tip can be measured from the deflection of the cantilever. The method can precisely control the force at a resolution of a few nN. The important points are a spring constant and a dashpot of the

^{a)}Electronic mail: hosaka@crl.hitachi.co.jp

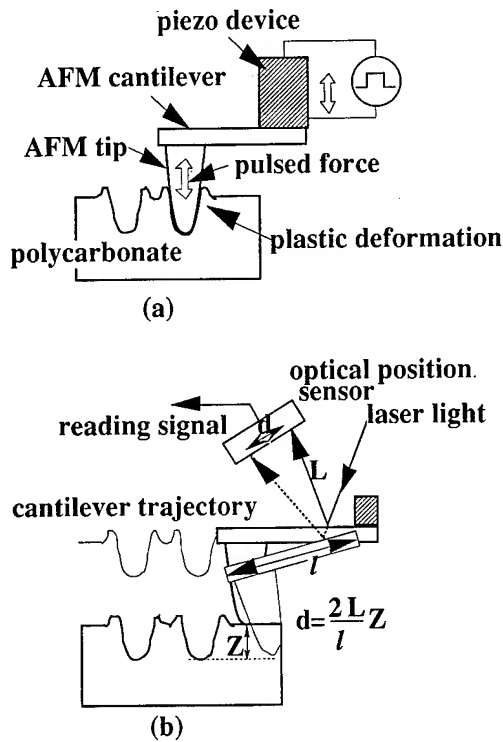


FIG. 1. Schematic diagram of new force modulation AFM pit recording; (a) nanopit formation (writing) and (b) readout of pits (reading).

sample surface in the model. When the spring constant and the dashpot are k^* and η , respectively, we can obtain the following equations:

$$m\ddot{X} = k(X_0 - X) - k^*X - \eta\dot{X}, \quad (1)$$

$$\frac{X}{X_0} = \frac{k}{(k^* + k) \left(1 + \frac{\eta}{(k^* + k)} S + \frac{m}{(k^* + k)} S^2 \right)} \quad (S = j\omega t). \quad (2)$$

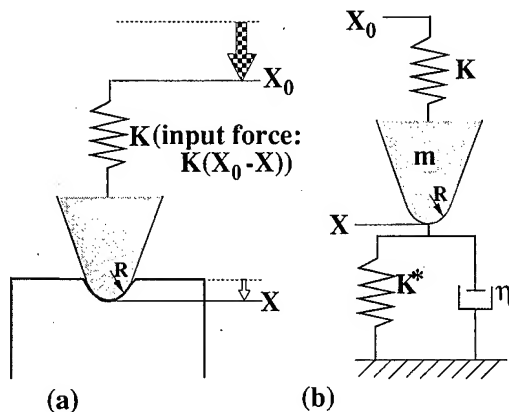


FIG. 2. Schematic diagram of pushing cantilever tip by external piezodriving and the tip penetration into sample surface; (a) simple model and (b) model equivalent to (a).

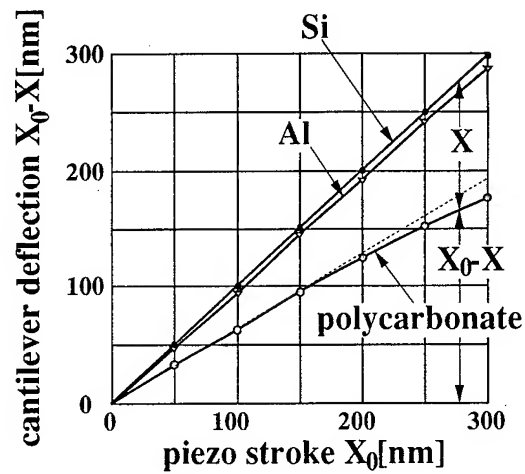


FIG. 3. Relationship between cantilever deflection and piezostroke.

Assuming that the time transient of the tip movement can be neglected at a frequency of less than 10 MHz ($1 \gg |\eta/(k^* + k)S + m/(k^* + k)S^2|$), the tip penetration depth can be given as follows:

$$X = \frac{k}{(k^* + k)} X_0, \quad X = \frac{k}{k^*} (X_0 - X). \quad (3)$$

Equation (3) shows that the tip penetration depth is in proportion to the tip working force and that it is in inverse ratio to a sum of spring constants of the tip and the sample surface in a region of elastic deformation.

The relationships between the cantilever deflection $X_0 - X$ and the piezostroke X_0 were measured as shown in Fig. 3. We performed the experiments in air using silicon(001), 1- μ m-thick aluminum on silicon and polycarbonate as the samples. The tip was silicon nitride. The cantilever deflection is the same as the piezostroke in the case of a silicon surface. On the other hand, the cantilever deflections of aluminum and polycarbonate gradually become less than the piezostroke. In particular, the cantilever deflection of polycarbonate is very much smaller than those of silicon and aluminum. This means that silicon, or aluminum, has a hard surface but polycarbonate has a soft surface. The tip penetration depths shown in Fig. 4 were obtained from these curves. The depth in the aluminum surface is less than 10 nm when the working force is less than 100 nN. In the polycarbonate surface, however, the tip penetrated into the surface to a depth of 40 nm at a working force of 40 nN. The elastic region may be up to 40 nN in the polycarbonate. The working force is gradually separated down with the tip penetration depth. This means that the critical force for plastic deformation was about 40 nN. In the elastic region, the working force is in proportion to the penetration depth. From this figure, the spring constant of the polycarbonate surface is about 1 N/m, and the relationship between the tip penetration and the working force follows Hooke's law, but not Hertz theory.

Figure 5 shows variations in the pit size on the polycarbonate surface due to a pulsed force for a pulse width of 6 ms. The pit size can be controlled by the pulsed force. From

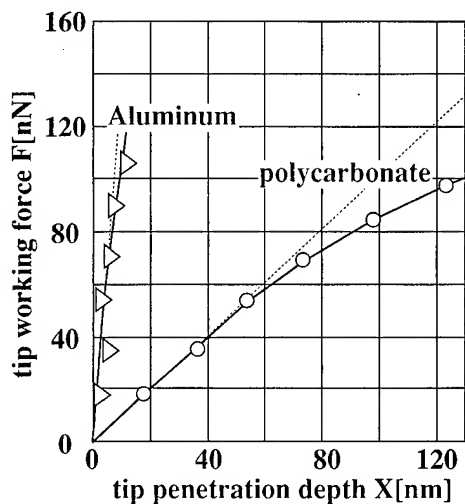


FIG. 4. Variations in tip working force due to tip penetration into sample surface, derived from Fig. 3.

the figure, the minimum size is less than 10 nm in diameter and the threshold force for plastic deformation is about 40 nN for 6 ms pulsed width. The threshold force agrees with the critical force in Fig. 4. The plastic deformation is considered to occur at a force of larger than 40 nN and it can be used for writing nanopits to achieve a trillion-bits-ultrahigh density recording.

IV. PROTOTYPE OF AFM RECORDING SYSTEM AND EXPERIMENTAL RESULTS

A. Prototyping of rotation type AFM recording system

We prototyped a rotation type AFM read/write experimental system. The system consisted of an AFM head system including a laser beam deflection detection system, a Z-axis fine and coarse actuator for approaching the tip to the sample surface, a controller, and an air spindle motor (Fig. 6). The revolution of the air spindle was controlled from 0.01 to 100 rpm within a deviation in speed of 10%.

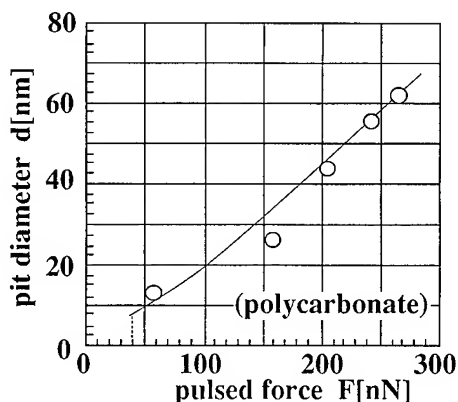


FIG. 5. Variations in pit diameter due to plastic formation created by pulsed force for pulsed width of 6 ms.

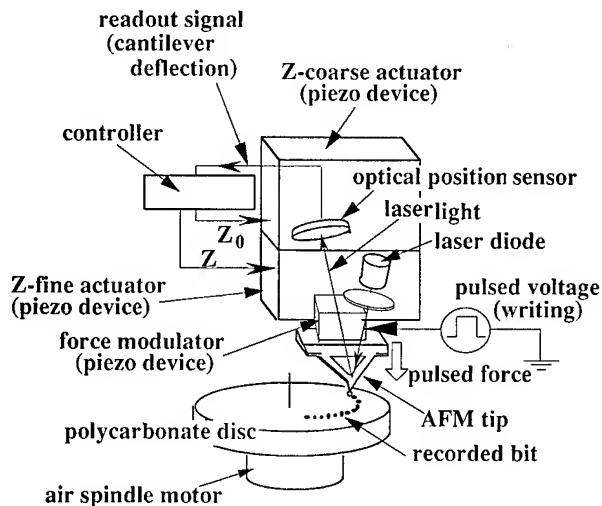


FIG. 6. Schematic diagram of rotation type AFM recording system.

B. High resonant frequency AFM cantilever

We developed a short length Si cantilever (Fig. 7). The lengths of 20 and 30 μm were prototyped. The thickness was about 0.3 μm . They are triangular and Λ type in shape. These cantilevers were fabricated by the Si microprocess. The process mainly consists of lever formation and tip formation. A silicon on insulator (SOI) wafer was used as the

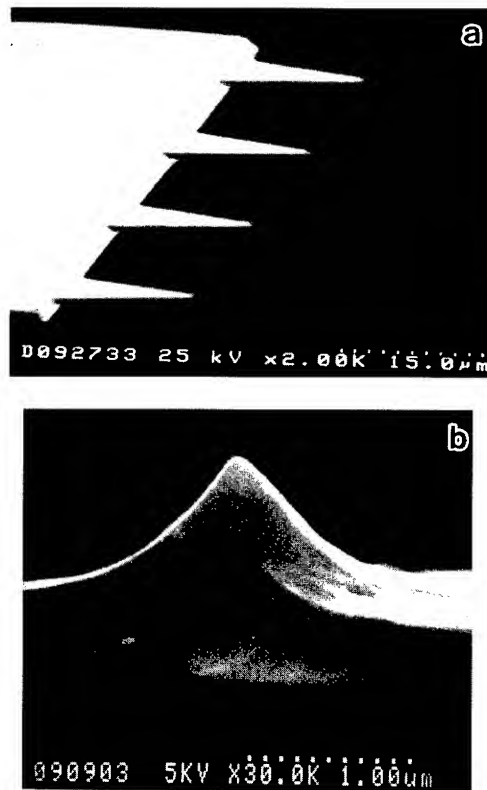


FIG. 7. SEM images of prototype of high resonant frequency AFM Si cantilever, (a) overall view and (b) tip detail.

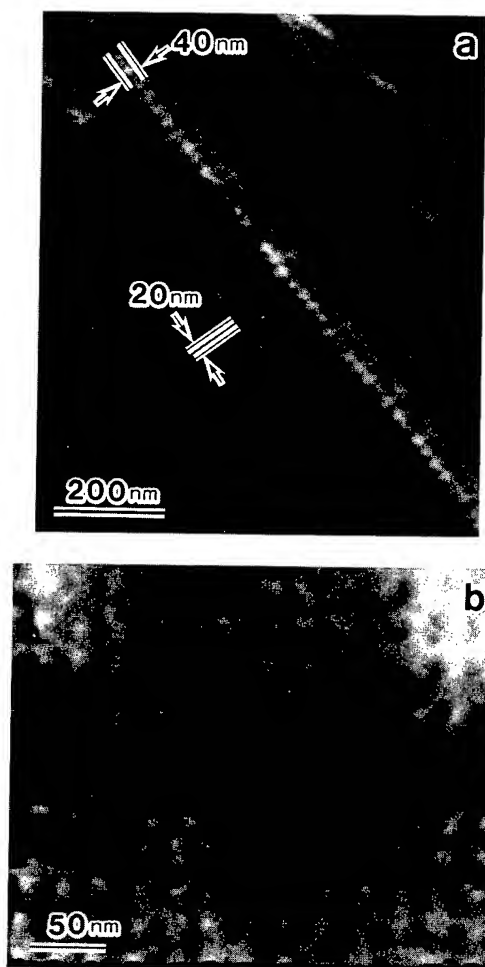


FIG. 8. AFM images of nanopit arrays, (a) formed by pulsed force of about 60 nN for 1 ms; which corresponds to 1.2 Tb/in.² in 1/2(2,7) code recording, and (b) by about 40 nN for 60 μ s.

substrate. The shape of the lever was made in the Si active layer by $\text{CF}_4 + \text{O}_2$ etching, potassium hydroxide (KOH) back-etching, etc. Then, the tip itself was mainly fabricated by SF_6 etching.

The first resonant frequency of the 30- μ m-length triangular type cantilever of 1.1 MHz was measured without contact. Then, the frequency was shifted to higher than 3 MHz in contact. The spring constant is estimated to be about 1 N/m.

C. Writing and reading

In the experiments, nanopit arrays with a 20 nm pitch in line and a 40 nm pitch in track were demonstrated (Fig. 8). The nanopit arrays were written at a pulsed force of 60 nN for 1 ms [Fig. 8(a)]. The array pattern means 1.2 Tb/in.² in the 1/2(2,7) code recording technique. Furthermore, a fine bit array with less than a 10 nm diam bit and a 25 nm pitch on the polycarbonate disk was formed [Fig. 8(b)]. These were made by applying a pulsed force of about 40 nN for 60 μ s to

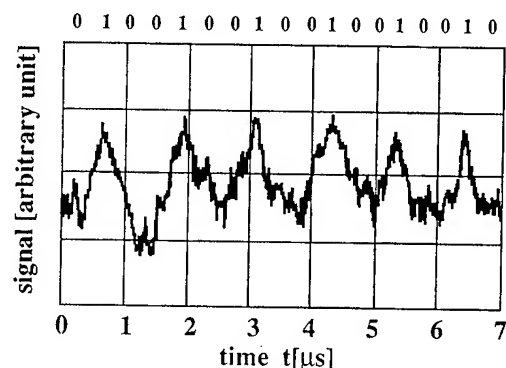


FIG. 9. Readout signal of a part of the nanopit array with a pit diameter of 25 nm at a spindle revolution of 100 rpm, which corresponds to 1.25 Mb/s in 1/2(2,7) code recording (preset force: 10 nN).

the tip. The nanopit arrays were recorded at a yield of 100%. We confirmed that the cantilever probe was able to write more than 10 million nanopits.

After the writing process, we read a part of the pit array at a readout speed of 1.25 Mb/s without tracking (Fig. 9). The rotation speed was 100 rpm and the constant force was set at about 10 nN. The minimum pit size was about 25 nm in diameter. As described above, AFM pit recording has the potential to achieve ultrahigh density recording with a density of over 1 Tb/in.² and a readout speed of over 1 Mb/s. Furthermore, a readout speed of over 10 Mb/s is considered possible by developing a cantilever with a higher resonant frequency.

V. SUMMARY

We proposed FM-AFM pit recording and demonstrated the possibility of achieving ultrahigh density recording with a high speed readout. We reached such a conclusion as our proposed FM-AFM pit recording is expected to be an attractive candidate for future ultrahigh density recording. And we achieved the following:

- (1) A minimum pit size of around 10 nm in diameter was formed by cold plastic deformation of polycarbonate disk surface.
- (2) The tip penetration depth into the polycarbonate surface was over 40 nm when the plastic deformed.
- (3) The nanopitting in the polycarbonate surface was formed at a force of over 40 nN.
- (4) A rotation type FM-AFM pit recording system was prototyped.
- (5) Using the prototyping, an ultrahigh recording density of 1.2 Tb/in.² and a readout speed of 1.25 Mb/s were demonstrated in 1/2(2,7) code recording.

However, there are still many subjects to be solved for data storage. We need to develop technologies in the future such as nanotracking, higher speed readout, nanotribology, etc.

ACKNOWLEDGMENTS

The authors wish to thank Dr. R. Imura and Dr. Y. Wada, T. Shintani and M. Muranishi of Hitachi, Ltd. for their fruitful discussion. We would also like to thank K. Matsumura, K. Muraoka, and T. Ohkawa of Hitachi, Ltd. for prototyping the rotation type AFM recording system. This work was supported by the Japan Key Technology Center.

¹T. R. Albrecht, M. M. Dovek, M. D. Kirk, C. A. Lang, C. F. Quate, and D. P. E. Smith, *Appl. Phys. Lett.* **55**, 1727 (1989).

²For example, S. Hosaka, S. Hosoki, T. Hasegawa, H. Koyanagi, A. Kikukawa, T. Shintani, and M. Miyamoto, *J. Vac. Sci. Technol. B* **13**, 2813 (1995).

³S. Hosaka, *IEEE Trans. Magn.* **32**, 1873 (1996).

⁴S. Hosaka, T. Shintani, M. Miyamoto, A. Kikukawa, A. Hirotsune, M. Terao, M. Yoshida, S. Honma, and S. Kammer, *J. Appl. Phys.* **79**, 8082 (1996).

⁵E. Betzig, J. K. Trautman, R. Wolfe, E. M. Gyorgy, P. L. Finn, M. H. Kryder, and C.-H. Chang, *Appl. Phys. Lett.* **61**, 142 (1992).

⁶H. J. Mamin and D. Rugar, *Appl. Phys. Lett.* **61**, 1003 (1992).

⁷S. Hosaka, H. Koyanagi, and A. Kikukawa, *Jpn. J. Appl. Phys.* **32**, L464 (1993).

⁸H. Kado and T. Tohda, *Appl. Phys. Lett.* **66**, 2961 (1995).

⁹R. C. Barrett and C. F. Quate, *J. Appl. Phys.* **70**, 2725 (1991).

¹⁰T. A. Jung, A. Moser, H. J. Hug, D. Brodbeck, R. Hofer, H. R. Hidber, and U. D. Schwartz, *Ultramicroscopy* **42-44**, 1446 (1992).

¹¹E. Hamada and R. Kaneko, *J. Phys. D* **25**, A53 (1992).

¹²S. Hosaka *et al.*, Japanese Patent No. P2533728 (submitted: 10 July 1987).

¹³G. Meyer and N. M. Amer, *Appl. Phys. Lett.* **53**, 1045 (1988).

Nanofabrication of electrodes with sub-5 nm spacing for transport experiments on single molecules and metal clusters

A. Bezryadin^{a)} and C. Dekker

Department of Applied Physics and DIMES, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands

(Received 20 March 1997; accepted 16 May 1997)

Electron-beam deposition is used to fabricate free-standing carbon nanoelectrodes separated by a gap of less than 5 nm. Fabrication is carried out under direct visual control in an electron microscope. After coating the carbon electrodes with a thin metal film (e.g., AuPd) such structures can be used to study electrical transport properties of single molecules (e.g., conjugated polymers) or metal nanoclusters. The fabrication process of the nanowires is described in detail. Furthermore, we suggest electrostatic trapping as a new method to bridge the electrodes with a single conducting nanoparticle in a controlled way. This principle was tested successfully on Pd nanoclusters and carbon nanotubes. © 1997 American Vacuum Society. [S0734-211X(97)13704-0]

I. INTRODUCTION

There has been much speculation about the prospect of "molecular electronics," i.e., electronic devices with single molecules as the active elements.¹ This may provide an attractive direction for the future development of electronic devices for several reasons. It conceptually can give the highest possible degree of circuit integration, viz., devices at the atomic scale. Devices may also operate at extremely high frequencies. Since different combinations of atoms give an almost infinite variety of chemical properties of resulting molecules, as is obvious from chemistry, one may expect that molecules with desired electronic properties can be synthesized as well. Many examples of molecules with an electronic functionality have already been synthesized, e.g., quantum wires such as conjugated polymers, molecular switches, molecular rectifiers, etc.^{1,2} Direct dc measurements of electrical properties of single molecules are lacking however. Usually experiments probe the conducting properties of a macroscopic amount of molecules. Here, hopping of electrons between molecules dominates transport. As a consequence, the intramolecular charge transport has not been well studied. This strongly contrasts the interesting theoretical predictions for electron transport through single molecular chains.³ Direct measurements of transport through single molecules will allow the verification of this, and are prerequisite for an assessment of the feasibility of molecular-scale electronics. Such experiments constitute the motivation for the technical developments of nanoelectrodes and electrostatic trapping reported in this article. The new techniques presented here may enable transport studies on single molecules.

There are two major problems which have to be solved to enable electrical transport experiments on single molecules. First, it is necessary to prepare conducting electrodes separated by a distance which at most equals the length of the molecule which has to be measured. This presents a difficulty because the available molecules of interest are usually

much smaller than the resolution limits (~ 10 – 20 nm) of the conventional electron beam lithography.⁴ For example, rigid conjugated oligomers of well-defined length are usually up to about 5 nm in length.⁵ Second, one has to find a way to bridge those electrodes with a single molecule. In this article we present a solution to both two obstacles.

Our method of fabrication of small gaps takes advantage of the electron beam deposition (EBD) of amorphous carbon which can be achieved in a scanning electron microscope (SEM).⁶ The SEM provides direct visual control which is an important advantage with respect to standard lithography. With EBD we can grow two free-standing carbon wires towards each other and adjust the gap between them to be as small as 3 nm. This is sufficiently small to bridge the electrodes by available conjugated oligomers.⁵

The principle of EBD is known since the pioneer work of Broers and co-workers.^{4,6} If the electron beam in a SEM is focused at a spot, it locally decomposes organic contamination molecules which usually are present in a small amount at the sample surface. It can, for example, be an ultrathin film of the pump oil due to a bad vacuum of the system. The decomposition of organic material leads to the deposition of a rigid⁷ solid material which is a mixture of amorphous carbon and some polymers.⁸ As soon as the contamination molecules in the e-beam spot are transformed into amorphous carbon, new organic material diffuses from other parts of the sample surface to the top of the carbon hill. The growth process thus proceeds continuously in time. As a result, by fixing the e-beam at a spot, a vertical carbon needle with a high aspect ratio can be deposited.⁸ By scanning the beam it is possible to locally cover the sample with carbon, which subsequently can be used as a protecting mask during an etching procedure. This method is known as contamination lithography⁶ which has a high resolution of about 10 nm.

Deposition of individual molecules or metal nanoclusters between two metallic electrodes is an issue which is being pursued by many research groups.^{9–11} In order to bridge the electrodes with a single conducting molecule or a metal cluster we suggest a new method, viz., electrostatic trapping

^{a)}Electronic mail: bezryadi@qt.tn.tudelft.nl

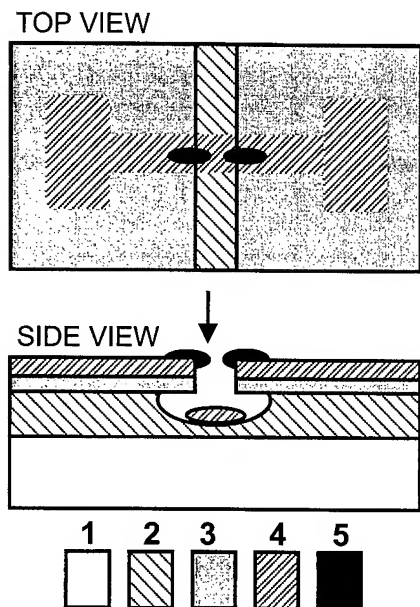


FIG. 1. Schematic view of the sample. It includes (1) a Si wafer, (2) a 1 μm thick SiO_2 film, (3) a low stress 60 nm thick film SiN, (4) a 10 nm thick AuPd film with contact pads, and (5) electron beam deposited carbon electrodes separated by a small gap, possibly less than 5 nm. The arrow shows the orientation of the electron beam during the growth of carbon nanoelectrodes. A metal film (not shown) is sputtered on top of the free-standing carbon electrodes to improve their conductivity.

(ET). This is a universal technique which can be used with any type of conducting particle. The principle is the following. Suppose we have two closely spaced electrodes which are immersed in a dilute solution of molecules or nanoclusters in some nonconducting solvent. If now a voltage of, say, 1 V is applied to the electrodes, then a very strong electric field, 10^8 V/m for a gap of 10 nm, will appear. Note that the electric field will exhibit a strong gradient, falling off quickly outside the gap. This field will polarize conducting nanoparticles in the neighborhood of the gap. The polarized particles will be attracted to the point of the strongest field, i.e., to the region right between the electrodes. The positively charged side of the molecule is attracted to the negative electrode and the negative side to the positive electrode. The electrodes thus will be connected by the molecule if the distance between them is sufficiently small.

Below we discuss the sample layout developed for experiments on single molecules (Sec. II), then we reiterate the principle of EBD and present details of fabrication of the free-standing carbon wires interrupted by a nanogap (Sec. III). Metal coating of these is presented in Sec. IV. Finally, we present some first results of electrostatic trapping of Pd nanoparticles and carbon nanotubes (Sec. V).

II. SAMPLE LAYOUT FOR TRANSPORT EXPERIMENTS ON SINGLE NANOPARTICLES

The sample layout is shown in Fig. 1. The sample consists of a silicon wafer covered with a film of SiO_2 and a thin low-stress SiN film. A narrow slit is etched in the SiN film in

SF_6 plasma. In the next step we slightly underetch the SiN film in HF acid to form an undercut. Due to this undercut, a thin AuPd film sputtered across the slit has an interruption of the order of 100 nm in width. Within this gap we grow two amorphous carbon needles towards each other. The needles are grown in a SEM using the electron beam deposition technique. An advantage of this layout is that the electrodes are free-standing, so one can evaporate metal on them (to improve their conductivity or to reduce the spacing between them) without making a short cut. The leakage resistance between the gold contact pads is very high, about $10^{13} \Omega$ due to the thick SiO_2 layer. Another advantage is the possibility to image trapped particles with SEM, or—after a slight modification such as etching a hole through the Si wafer—with transmission electron microscopy (TEM).

The fabrication steps are as follows. We start with a Si wafer covered with a 1 μm thick thermally grown siliconoxide layer, and a 60 nm thick ultralow-stress SiN film made by low-pressure chemical-vapor deposition.¹² The next step is the etching of a long slit in the SiN film. Before etching, a 300 nm film of poly(methylmethacrylate) (PMMA) (950 K) resist is spun at 5000 rpm onto the sample to protect the SiN film. Patterning of the resist is carried out in an electron beam pattern generator. The slit usually is about 100–150 nm in width and a few mm in length. After the slit is made by means of reactive ion etching in a SF_6 plasma, we slightly underetch the underlying SiO_2 with HF acid. The etching rate of SiN in 40% HF solution is 4 nm/min; SiO_2 is etched at a rate of about 1500 nm/min. Underetching during about 20 s results in a free-standing SiN membrane of approximately 500 nm in width divided into two parts by the slit as it is shown in Fig. 1. Next we sputter across the slit a strip of gold or gold palladium of 10–20 nm in thickness and about 20 μm in width. For the pattern definition of the strip a mechanical shadow mask (not shown) is used. The mask is made from a Si wafer (100) by anisotropic etching in KOH. The slit cuts the gold strip into two parts separated by the 100 nm gap. It is inside this gap that we grow the carbon nanowires starting from the edges of the gold film. Due to the underlying SiO_2 layer the leakage resistance between the Au electrodes is about $10^{13} \Omega$ while the total area of Au electrodes including the contact pads is a few square millimeters. The high value of the leakage resistance is important, because it acts as a parallel resistance to that of the single nanoparticle which is localized between the carbon electrodes during the measurements. And we will see below, a single particle can have a rather low conductivity.^{13,14}

III. ELECTRON BEAM DEPOSITION OF FREE-STANDING CARBON NANOWIRES

The last two steps of the fabrication process are the growth of two carbon nanowires towards each other and coating them with a thin metal film to improve their conductivity. These issues will be described in details in this and the next section, respectively. As indicated in Fig. 1, the nanowires are attached to the gold electrodes with the underlying SiN membrane and stay within the plane of the membrane.

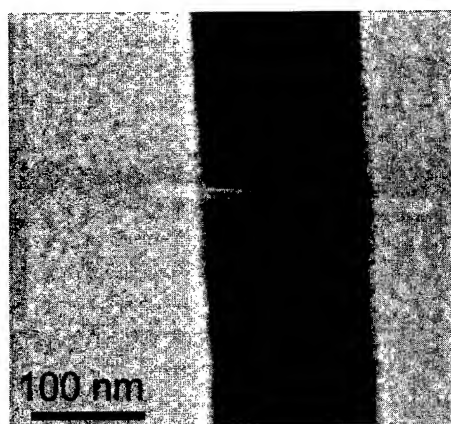


FIG. 2. A free-standing amorphous carbon wire of about 10 nm in width, grown by a slow linear scan of the electron beam across the slit in the SiN membrane. The electron beam deposition was done after sputtering of a 20 nm thick AuPd film on the membrane. The metal film makes electrical connection to the carbon wire and provides a convenient amount of organic contamination which is necessary for the EBD. The electron beam scan rate was about 3 nm/s.

The growth process should be stopped when the desired distance between their tips is achieved. We have succeeded in making this distance as small as 3 nm. The low limit is determined by the resolution of SEM (about 1 nm for our Hitachi S-900 SEM) and the growth rate which is between 1 and 10 nm/s in our case. The latter can be reduced if necessary by cleaning the sample in oxygen plasma or stopped by cooling to 77 K.

In practice, free-standing horizontal carbon nanowires are fabricated in the following way. We focus the beam on the gold film near the edge of the membrane and then move it slowly (~ 1 nm/s) towards the slit. Initially the deposition occurs at the gold film but as soon as the beam leaves the membrane a free-standing carbon needle starts to grow parallel to the membrane, inside the slit. If the e-beam sweep rate is too high, i.e., higher than the deposition rate, the needle cannot follow the beam and the growth process is interrupted. If on the other hand the e-beam velocity is too slow, a sheet of carbon rather than a rod is formed with a large size in the direction of the beam, which is much higher than the width in the direction perpendicular to the beam and to the motion. To grow a thin symmetric carbon rod we start scanning the beam with a high velocity across the slit periodically, and then slowly decrease the beam velocity until the growing carbon wire can follow the beam. The width of electron beam deposited needles usually is between 5 and 20 nm (see for example, Fig. 2). If necessary it can be reduced down to about 3 nm by a slow etching in an oxygen plasma. The electron beam deposited carbon is poorly electrically conducting, especially at low temperatures. Consequently carbon nanowires can be used as a template for the fabrication of narrow one-dimensional metal wires, or more sophisticated patterns by simply depositing a metal film on top.¹⁵

An alternative way of making free-standing structures takes advantage of the feedback control of the signal of secondary electrons.¹⁵ In this case the beam is positioned *in* the

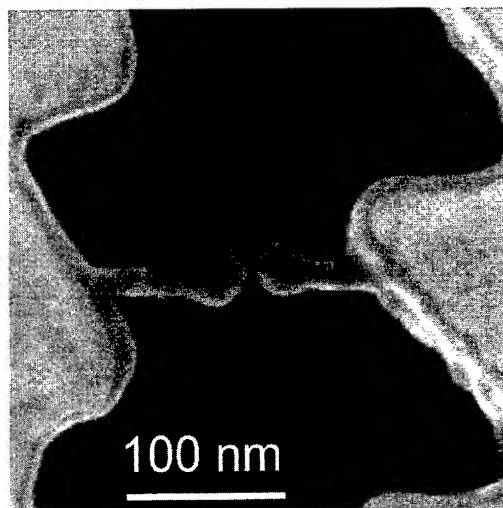


FIG. 3. Two free-standing carbon nanowires grown towards each other using the electron beam deposition technique. They are separated by a distance of about 4 nm. Gold grains visible behind the whiskers are not in focus because they are lying about $0.5 \mu\text{m}$ below the nanowires, on the bottom of the groove (see the sample layout in Fig. 1). The small value of the gap is achieved by zooming in the SEM on the area around the gap. This results in a thickening of the tips of the electrodes. The acceleration voltage and the e-beam current were 30 kV and about 20 pA correspondingly.

slit, close (~ 5 nm) to the membrane edge. Initially the signal of secondary electrons is zero because the beam does not cross anything. The deposition of carbon starts on the membrane regardless of the fact that the beam does not touch it.¹⁶ After some time (~ 10 s) the carbon needle reaches the e-beam and some secondary-electron signal appears. When the signal reaches a certain level, we step the beam away from the membrane edge and wait again till the needle reaches the beam. This procedure is carried out repeatedly. Of course such an algorithm can also be applied to a continuously moving beam. The feedback process usually results in somewhat thicker rods than the linear scan method.

Now let us see how a gap of a few nm can be achieved. Of course one can grow two whiskers towards each other from the opposite sides of the slit. But it is difficult to make a small, say < 5 nm, gap from the very beginning, because it is not possible to observe the first nanowire while growing the second one, since the beam does not scan the area for imaging but moves along with the second carbon wire. However, an initial gap of, say, 30 nm can easily be made in this way. Then to reduce the gap it is sufficient to zoom in the SEM on the nanowires near the gap, i.e., the tips of both carbon needles should be visible on the SEM screen. During the ordinary imaging the process of continuous deposition of carbon slowly proceeds at all points of the area that is scanned by the electron beam. Consequently the nanorods grow in size and the gap between them decreases continuously. When the distance between their tips reaches desired value, the SEM is switched off. The smallest gap we did fabricate by this method was less than 3 nm. An example is shown in Fig. 3. Note that the amorphous carbon nanowires are very rigid and chemically inert, so the gap is very stable.

It does for instance not change when the sample is cleaned with different solvents or acids.

As we have seen, EBD is a powerful technique for nanofabrication. It provides the possibility of fabricating structures of nm dimensions under direct visual control. It does not require resist and, accordingly, EBD is not limited to fabrication of structures on a planar surface. It gives a possibility to grow free-standing nanostructures. A possible drawback is the use of a surface contamination as a precursor which commonly is supposed to arise from the bad vacuum. This may render the process not well-reproducible because the contamination depends on the cleanliness of the SEM machine. The level of the contamination in our Hitachi SEM is very low. Therefore the deposition is very slow or it does not occur at all if the sample has been cleaned by an oxygen plasma or through a wet etch in aqua regia ($\text{HNO}_3:\text{HCl}$, 1:3).

We have optimized the technique for better reproducibility by investigating cleaning procedures and well-characterized precursor materials. We have tried various organic liquids as precursors. The liquid should wet the surface to form a thin uniform film and it should not be too volatile in order to have it stay on the surface when the sample is exposed to the vacuum inside the SEM. It was found that liquid paraffin (a mixture of hydrocarbons from $\text{C}_{12}\text{H}_{26}$ to $\text{C}_{18}\text{H}_{38}$) gives a very high deposition rate of about 100 nm/s or higher. To form a thin paraffin film on the surface the sample can be immersed in a solution of the paraffin in cyclohexane (1:100) and then dried in a flow of nitrogen gas. It is a good precursor but the EBD rate is very high. Hexadecane ($\text{C}_{16}\text{H}_{34}$) turned out to be a more suitable precursor for our applications. This organic liquid wets our sample surface and can be decomposed by the electron beam. The presence of traces of this slightly volatile liquid on the surface provides a convenient EBD rate (from 1 to 10 nm/s). To cover the sample with this liquid precursor it is sufficient to put a drop of $\text{C}_{16}\text{H}_{34}$ on the surface and then dry the sample in a nitrogen flow during a few seconds. A convenient way to contaminate the surface and enable EBD is to sputter a thin metal film. In this case the contamination is due to the bad vacuum in the sputtering machine. In both cases ($\text{C}_{16}\text{H}_{34}$ precursor or sputtering contamination) EBD occurs in a very similar way and results in carbon wires of similar sizes and quality. EBD occurs also in the TEM under the influence of the parallel, very broad beam of electrons. By cooling the sample down to the liquid nitrogen temperature it is possible to stop the deposition process and image the sample without modifying the amorphous carbon nanostructures.

IV. METAL COATING OF THE CARBON ELECTRODES

The resistivity of EBD amorphous carbon usually is quite high (10^2 to $10^5 \Omega \text{ cm}$ at room temperature) and rather irreproducible. For transport experiments one needs electrodes with a sufficiently high and predictable conductivity. This can be achieved by coating the carbon needles with an amorphous or small-grain-size metal film, by sputtering or evaporation. If the experiment has to be done in air the coating

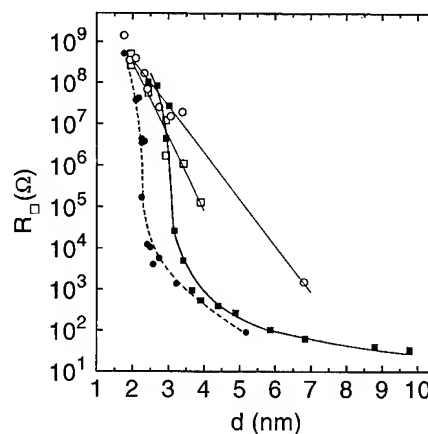


FIG. 4. Square resistance R_{\square} vs film thickness, d , for various sputtered thin films. Solid squares represent AuPd (4:1) films on SiN. Solid circles denote AuPd films on a very thin (0.3 nm) sticking layer of chromium which serves to suppress the granularity. Here R_{\square} is shown vs the total thickness of AuPd and Cr film. Open symbols show the square resistance of films which are sputtered onto carbon wires grown by EBD.

material should not oxidize. Also the metal film should not be thick, otherwise it will reduce or even close the gap between the electrodes. We have investigated the efficiency of the gold palladium (4:1) as a coating material. The square resistance of films sputtered on a smooth dielectric surface (SiO_2 or SiN) is plotted versus the film thickness in Fig. 4 (solid squares). The resistance starts to rise very sharply,¹⁷ even on this log scale, when the thickness becomes lower than 3 nm. SEM imaging shows that below this value the film consists of isolated grains. The granularity can be reduced if one first sputters a few Å of Cr. In this case, even a 2.5 nm film has already a quite low resistance, $R_{\square} \sim 10^4 \Omega$ (Fig. 4, solid circles). Note that after sputtering of the Cr film, the sample was exposed to air for about 1 min before the subsequent sputtering of AuPd was carried out.

A qualitatively different behavior of the film resistance versus thickness is found when the film is sputtered onto the carbon wires such as shown in Fig. 2. The resistance of such films drops exponentially with decreasing thickness, which is slower than for films on a planar substrate. The difference is probably due to the fact that in two dimensions (on the planar substrate) a percolation path appears at some critical size of the grains. This leads to the jumpwise drop of the square resistance. The metal film on a carbon nanowire is effectively one-dimensional because the metal-grain size is of the same order of magnitude as the wire diameter. Therefore there is only one possible path, and it is improbable that all grains on this path are touching each other. Now if there is at least one interruption, then electrons have to tunnel across it (probably through the carbon substrate). The size of interruptions will probably be inversely proportional to the film thickness. The resistance then is an exponential function of the film thickness. The current-voltage characteristics (not shown) of all ultrathin coating films are linear.

The metal coating technique also provides an alternative way of making small gaps. In Fig. 5 we present two nan-

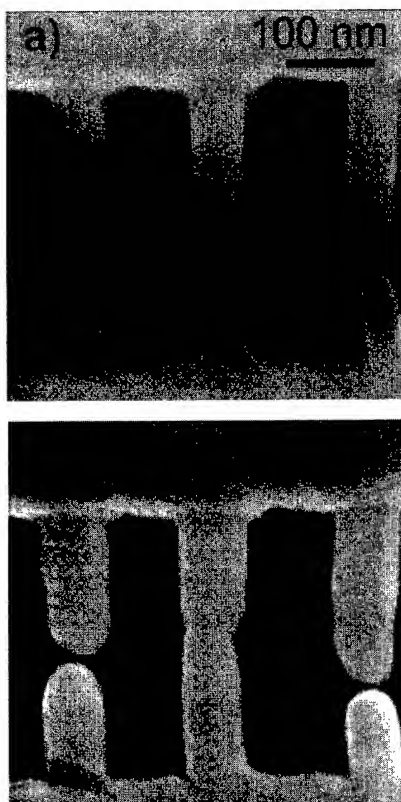


FIG. 5. Free-standing carbon nanowires (bright in the pictures) (a) before and (b) after coating with a thin AuPd film. The coating reduces the spacing between the electrodes by approximately 11 nm (from 16 to 5 nm). Note that this is a factor of 3 smaller than the sputtered film thickness which is about 30 nm in the present case.

ogaps between carbon needles, before and after coating with a 30 nm film of AuPd. By choosing the appropriate size of the initial gap and the AuPd film thickness, it is possible to reduce the gap down to 5 nm by coating, as shown in Fig. 5. Note that right after the coating the sample was cleaned in oxygen plasma (50 W) for about 5 min. In independent tests we have verified that such cleaning does not etch the AuPd electrodes. The cleaning is necessary to remove organic contamination to enable the high-resolution imaging without depositing a new layer of carbon. It was also found that the oxygen plasma etches away the underlying carbon so one can obtain free-standing gold palladium wires.

We thus have demonstrated two ways of making electrodes separated by a very small (<5 nm) distance. In the first technique, a nanogap of *desirable* size is fabricated by EBD under visual control in SEM. Then electrodes are coated with a few nm (say 4 nm) thin film to improve their conductivity. The gap reduction caused by the coating is small, about 1–2 nm in this case, because, as shown in Fig. 5, the gap reduction is smaller than the sputtered film thickness by a factor of ~ 3 . The second possibility is to start with a quite large (e.g., 20 nm) initial gap and then sputter sufficient metal to reduce the gap size to the desired value. The smooth coating [Fig. 5(b)] makes this process quite reproducible, but nevertheless a few steps of sputtering usually are

necessary before the desired small gap size will be achieved. Let us mention also a third possibility, which has not been verified in our study. By using organometallic compounds as precursor material one can grow needles which consist of metallic nanograins surrounded by amorphous carbon.¹⁸ This may yield a rather low resistivity ($0.01 \Omega \text{ cm}$). This implies a carbon electrode resistance of $\sim 10 \text{ k}\Omega$ (without coating) which is sufficiently low for molecular transport experiments.

V. ELECTROSTATIC TRAPPING OF SINGLE CONDUCTING NANOPARTICLES

Suppose we have two electrodes with a spacing smaller than the length of a molecule or a metal cluster. Then, a single nanoparticle may be fixed between the electrodes by means of ET. The principle of ET is the following. First we apply a voltage to the electrodes thus creating a strong electric field in the gap. If a conducting nanoparticle comes close to the gap it will be polarized by the electric field near the gap. Then, like any dipole, it will be oriented along the field, and will be attracted to the region of the strongest field, i.e., to the gap between electrodes. The positive side of the polarized molecule will be attracted to the negative electrode and the negative side to the positive. The molecule thus will bridge the electrodes. This principle of electrostatic trapping is expected to work for any type of conducting nanoparticles, i.e., molecules with delocalized electrons, metallic clusters, carbon nanotubes, etc.

To bring the molecules close to the electrodes one by one, one can immerse the electrodes in a dilute solution of molecules. Another possibility would be to expose the electrodes to a vapor of molecules in an ambient inert gas. Molecules will diffuse and from time to time come close to the gap. Here they will be trapped electrostatically as explained above. As soon as one molecule is trapped, a current will flow between the electrodes, and the sample can be withdrawn from the solution or vapor. Clearly, a good solvent for ET should not conduct electrically, so there is no current between the electrodes before a molecule is trapped. Also it should not be polar, since this would lead to screening of the field in the gap.

It is interesting to note that if one applies a series resistor R_s between the voltage source and electrodes, with a value much higher than the resistance of the molecule R_m , then the electrostatic field in the gap will be strongly reduced (by a factor R_s/R_m) as soon as one molecule is trapped. As a consequence, trapping of a second particle may be prevented.

We have tested ET on carbon nanotubes and palladium colloid particles. These have been chosen because of their relatively big size which provides the possibility to image them with SEM and directly verify the efficiency of the ET. Carbon nanotubes are long cylindrical molecules containing only carbon atoms.¹⁹ Single-wall nanotubes were synthesized by Smalley and co-workers.²⁰ They are not solvable in any known solvent but can be ultrasonically dispersed in various organic liquids. We have used cyclohexane where a small amount of soot containing the nanotubes was dispersed

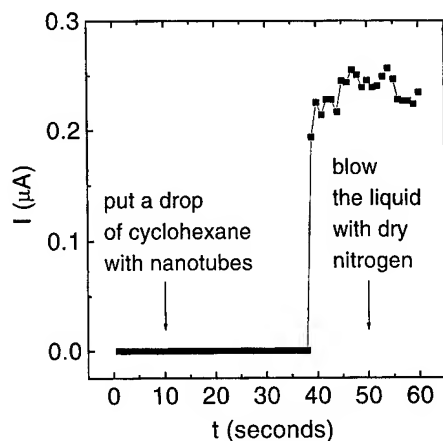


FIG. 6. Variation of the current flowing between the electrodes during a trapping experiment. The voltage was fixed at 4.5 V; the series resistor was 10 M Ω . The distance between the electrodes was about 150 nm.

ultrasonically. After excitation during a few minutes the soot was not visible any more and the suspension was transparent. It contains single nanotubes, bundles of parallel nanotubes ("ropes"²⁰), tangles of ropes, and pieces of amorphous carbon. Therefore only a part of the trapping experiments resulted in a single rope lying across the electrodes.

Trapping events can be detected by measuring the current between the electrodes in real time. During ET the current usually changes as shown in Fig. 6. In this experiment we applied 4.5 V to the electrodes through a series resistor of 10 M Ω . Then we put a drop of the nanotube dispersion in cyclohexane on the electrodes at $t = 10$ s. After about 30 s a finite current was detected which indicates that a nanotube rope is trapped on the electrodes. Upon drying the liquid at $t = 50$ s the current did not change. In some cases a considerable reduction (about one order of magnitude) of the resistance was observed during the drying. Examples of nanotube ropes that are trapped electrostatically are shown in Fig. 7. In Fig. 7(b) we observe a toroidal nanotube. This circular form of carbon nanotubes was discovered recently.²¹ Transport experiments on the nanotube ropes will be reported elsewhere.

The ET method was also tested on Pd colloid particles which were synthesized by Schmid and co-workers.²² These are almost identical Pd nanoparticles of about 20 nm in diameter. They are covered with a layer of stabilizing ligands of about 1 nm in thickness which prevent the coalescence of the particles. Due to the organic shell, the particles are solvable in water. Water has a finite conductivity, which in fact is much higher than the conductivity of a single Pd particle bridging the electrodes. This renders it impossible to observe trapping events directly in real time. Trapping can be achieved, however, in the following way. First we put on the electrodes a drop of water with dissolved Pd particles, then apply a voltage of 4.5 V to the electrodes through a series resistor of 100 M Ω which leads to a current of about 30 nA. After a few seconds the voltage is reduced to 0.45 V, and the water is removed with a flow of dry nitrogen gas. The resistance between the electrodes before the trapping

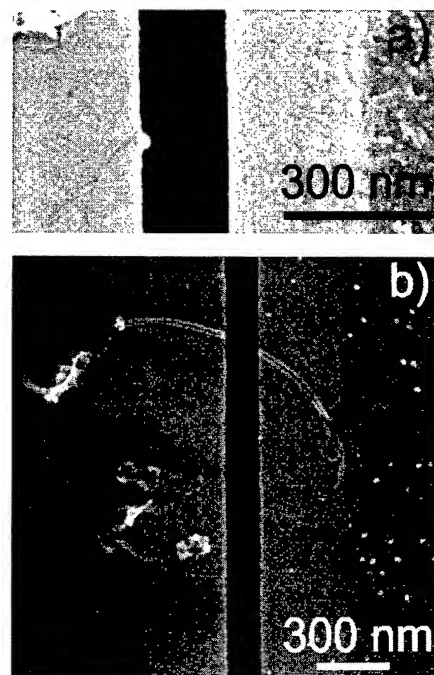


FIG. 7. (a) Bundle of nanotubes that has been trapped electrostatically between two AuPd electrodes separated by a slit of about 150 nm. (b) A toroidal nanotube (Ref. 21). The thicknesses of the ropes are about 6 and 25 nm, respectively.

procedure is very high (10^{13} Ω) and stays high after the drying if no particles were solved in the water, or if the voltage was not applied. Only if we use water with solved particles and apply a sufficiently high voltage, we observe a reduced resistance (a few G Ω) after the drying. Inspection with SEM shows that one or a few particles are trapped in the gap, depending on the distance between the electrodes. In Fig. 8 we present two examples.

Preliminary transport measurements show that the room temperature resistance of single Pd nanoparticles trapped between two electrodes has an exponential dependence on the bias voltage, i.e., $R \propto \exp(-V/V_0)$. The origin of this exponential behavior may tentatively be attributed to the voltage dependence of the tunnel barrier separating the particles from the electrodes. The barrier in this case is formed by the organic shell of the Pd particles. Tunneling through a monolayer of insulating organic molecules was studied by Boudas *et al.*,²³ who also observed an exponential drop of the resistance with increasing voltage. At 4 K we observe a gap in the current versus voltage characteristics indicating Coulomb blockade²⁴ effects. Details of the transport experiments will be reported elsewhere.

VI. CONCLUSIONS

We have demonstrated a technique of nanofabrication of electrodes separated by a gap of less than 5 nm. It is based on electron beam deposition of amorphous carbon and is carried out in a scanning electron microscope under direct visual control. Also we have suggested a new method of electrostatic trapping which provides a possibility to bridge the

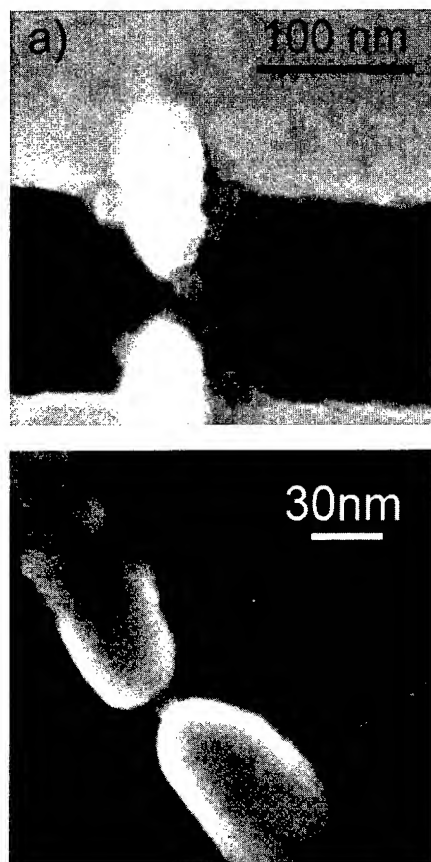


FIG. 8. Pd nanoparticles that are trapped electrostatically between electrodes. If the distance between the electrodes is smaller than the particle size then only a single particle is trapped as in (b). In the case of larger distances the electrodes are usually bridged with two or more particles as shown in (a). The metal films used for coating the electrodes were about 30 and 12 nm thick, respectively.

electrodes with a *single* conducting nanoparticle in a controlled way. This has been tested successfully for carbon nanotubes and Pd colloid particles. Transport measurements of single Pd colloid particles show an exponential dependence of the resistance on the applied voltage. Experiments on single conjugated molecules are in progress now.

ACKNOWLEDGMENTS

The authors are grateful to R. E. Smalley and G. Schmid for supplying carbon nanotubes and the Pd colloid, respectively. Interest and help of Y. Volokitin were important in the development of ET. Useful discussions with L. de Jongh, E. W. J. M. van der Drift, E. J. G. Goudena, P. M. Sarro, S. J. Tans, N. Gribov, and Y. Volokitin are acknowledged. This research was partly funded by the "Stichting voor Fundamenteel Onderzoek der Materie (FOM)."

- ¹A. Aviram and M. A. Ratner, *Chem. Phys. Lett.* **29**, 277 (1974); *Molecular Electronics—Science and Technology*, AIP Conf. Proc. 262, edited by A. Aviram (AIP, New York, 1992); *Introduction to Molecular Electronics*, edited by M. C. Petty, M. R. Bryce, and D. Bloor (Arnold, London, 1995).
- ²*Molecular Electronic Devices*, edited by F. L. Carter (Dekker, New York, 1982); L. Jones, D. L. Pearson, J. S. Schumm, and J. M. Tour, *Pure Appl. Chem.* **68**, 145 (1996).
- ³W. P. Su, J. R. Schrieffer, and A. J. Heeger, *Phys. Rev. Lett.* **42**, 1698 (1979); C. Joachim and J. F. Vinuesa, *Europhys. Lett.* **33**, 635 (1996); M. P. Samanta, W. Tian, and S. Datta, *Phys. Rev. B* **53**, 7626 (1996); V. Mujica, M. Kemp, and M. A. Ratner, *J. Chem. Phys.* **101**, 6849 (1994).
- ⁴A. N. Broers, *IBM J. Res. Dev.* **32**, 502 (1988).
- ⁵P. Bäuerle, T. Fischer, B. Bidlingmeier, A. Stabel, and J. P. Rabe, *Angew. Chem. Int. Ed. Engl.* **34**, 303 (1995); D. M. de Leeuw, *Synth. Met.* **55–57**, 3597 (1993).
- ⁶A. N. Broers, W. W. Molzen, J. J. Cuomo, and N. D. Wittels, *Appl. Phys. Lett.* **29**, 596 (1976).
- ⁷M. Wendel, H. Lorenz, and J. P. Kotthaus, *Appl. Phys. Lett.* **67**, 3732 (1995).
- ⁸Y. Akama, E. Nishimura, A. Sakai, and M. Murakami, *J. Vac. Sci. Technol. A* **8**, 429 (1990).
- ⁹D. L. Klein and P. L. McEuen, *Appl. Phys. Lett.* **68**, 1 (1996).
- ¹⁰V. Rousset, C. Joachim, S. Itoua, B. Rousset, and N. Fabre, *J. Phys. III* **5**, 1983 (1995).
- ¹¹M. A. Reed, C. Zhou, C. J. Muller, T. P. Burgin, and J. M. Tour (unpublished).
- ¹²H. W. van Zeijl, L. K. Nanver, and P. J. French, *Electron. Lett.* **31**, 927 (1995).
- ¹³L. A. Bumm, J. J. Arnold, M. T. Cygan, T. D. Dunbar, T. P. Burgin, L. Jones II, D. L. Allara, J. M. Tour, and P. S. Weiss, *Science* **271**, 1705 (1996).
- ¹⁴H. van Kempen, J. G. A. Dubois, J. W. Gerritsen, and G. Schmid, *Physica B* **204**, 51 (1995).
- ¹⁵A. Yu. Kasumov, I. I. Khodos, N. A. Kislov, O. V. Kononenko, V. N. Matveev, V. A. Tulin, Yu. B. Gorbatov, V. I. Nikolaichik, and E. E. Vdovin, *Phys. Rev. Lett.* **75**, 4286 (1995); V. V. Aristov, A. Yu. Kasumov, N. A. Kislov, O. V. Kononenko, V. N. Matveev, V. A. Tulin, I. I. Khodos, Yu. A. Gorbatov, and V. I. Nikolaichik, *Nanotechnology* **6**, 35 (1995).
- ¹⁶V. V. Aristov, N. A. Kislov, and I. I. Khodos, *Microsc. Microanal. Microstruct.* **3**, 313 (1992).
- ¹⁷C. E. Ells and G. D. Scott, *J. Appl. Phys.* **23**, 31 (1952); U. Oppenheim and J. H. Jaffe, *ibid.* **24**, 1521 (1953).
- ¹⁸H. W. Koops, A. Kaya, and M. Weber, *J. Vac. Sci. Technol. B* **13**, 2400 (1995); T. J. Stark, Y. M. Mayer, D. P. Griffis, and P. E. Russell *ibid.* **10**, 2685 (1992).
- ¹⁹S. Iijima, *Nature (London)* **354**, 56 (1991).
- ²⁰A. Thess, R. Lee, P. Nikolaev, H. Dai, P. Petit, J. Robert, C. Xu, Y. H. Lee, S. G. Kim, A. G. Rinzler, D. T. Colbert, G. E. Scuseria, D. Tommanek, J. E. Fischer, and R. E. Smalley, *Science* **273**, 483 (1996).
- ²¹J. Liu, H. Dai, J. H. Hafner, D. T. Colbert, R. E. Smalley, S. J. Tans, and C. Dekker *Nature (London)* **385**, 781 (1997).
- ²²G. Schmid, *Clusters and Colloids. From Theory to Applications* (VCH, Weinheim, 1994); *Physics and Chemistry of Metal Cluster Compounds*, edited by L. de Jongh (Kluwer, Dordrecht, 1994).
- ²³C. Boulas, J. V. Davidovits, F. Rondelez, and D. Vuillaume, *Phys. Rev. Lett.* **76**, 4797 (1996).
- ²⁴*Single-Charge Tunneling*, edited by H. Grabert and M. H. Devoret (Plenum, New York, 1992).

Deconvolution of tip affected atomic force microscope images and comparison to Rutherford backscattering spectrometry

M. F. Tabet^{a)} and F. K. Urban III

Electrical and Computer Engineering Department, Florida International University, Miami, Florida 33199

(Received 14 April 1997; accepted 18 April 1997)

There is distortion in atomic force microscope (AFM) images caused by a nonideal shape and size of the probe tip. This is called tip effect and is due to the convolution of the tip and sample. AFM images of nanometer size islands of zinc deposited on silicon substrates by the ionized cluster beam deposition technique were used to investigate this effect. The number of zinc atoms per unit area determined by each of two methods, the AFM images and Rutherford backscattering spectrometry (RBS) are compared and the AFM reported more zinc than RBS. A partial explanation for this difference is that the convolution of the tip and sample makes the islands appear larger in the AFM data. Previously reported convolution and deconvolution algorithms were implemented to study and simulate the interaction between tip and sample in the AFM. The deconvolution algorithm removes part of the distortion by taking into account the physical volume occupied by the tip which exposes a more accurate image. After deconvolution of the zinc islands images there was better agreement between AFM and RBS results. Deconvolution of other images will also be discussed. © 1997 American Vacuum Society. [S0734-211X(97)06304-X]

I. INTRODUCTION TO IONIZED CLUSTER BEAM

The ionized cluster beam (ICB) technique is a thin film deposition method in which a beam of ionized and neutral atomic clusters of varying sizes is formed and deposited on substrates to form thin films. In 1972, Takagi *et al.* from Kyoto University in Japan, originated the idea of the ICB thin film deposition technique in which it was proposed that free jet clusters formed by high temperature evaporation of room temperature solids, would be used to form thin films.¹ The beam also passes through an ionizer where a fraction of the clusters are ionized and accelerated to the substrate along with the neutral beam for enhanced thin film formation. They postulated that clusters would enhance film nucleation and surface mobility, and film to substrate adhesion resulting in higher quality films. The researchers in Kyoto believed that the material escaping through the nozzle underwent adiabatic expansion and condensed into large clusters up to thousands of atoms in size.

Although Takagi *et al.* reported singly ionized clusters containing 10^2 – 10^3 atoms,² significant numbers of such large clusters were not reported for a range of materials by other researchers using the same conditions.^{3–5} The consensus now is that there was no significant number of large clusters in those earlier reports and the quality of the films was almost certainly due to accelerated ions in the depositing beam.^{3,4} Therefore, the effect of large-size clusters on film growth and characteristics is not clear.

Since 1992 large clusters of zinc, formed from a pure expansion of zinc vapor, in the range of 2000 atoms have been produced in our laboratory.^{6–8} Two major changes were made to the original cluster source, crucible pressure was increased from 2 to several thousand Torr and a converging-diverging nozzle 18 mm long and 0.4 mm in diameter at the

throat was used instead of the 1×1 mm nozzle used in the earlier work.

Urban *et al.* also developed the deflected deposition method in order to study the cluster size distribution in the beams.⁸ The ionized particles are deflected from a straight path to the substrate by a deflection field placed between the two deflection plates that are perpendicular to the beam direction. After passing through the deflection region, the ionized part of the beam is displaced and deposited on the substrate according to the energy of the particles. Assuming that each cluster is singly charged, the distribution of the cluster size is obtained from the ratio of the atom to charge flux as a function of deflection distance. The atom flux is calculated from data obtained by Rutherford backscattering spectrometry (RBS) along the substrate in the deflection direction, shown in Fig. 1. The charge flux is obtained from ion current measurements on substrate wires. The average cluster size was a function of crucible temperature, decreasing with a decrease in temperature. Average sizes of 100, 1000, and 2200 atoms/cluster were obtained at crucible temperatures of 1120, 1180, and 1260 °C.

Atomic force microscope (AFM) images were taken in the deflection region at selected points along a path parallel to the RBS scan line. These 3×3 μm images show a flat silicon substrate with zinc islands on it, whose number depends on the position of the scan on the substrate. Figure 2 shows an example of one of these images. The AFM is capable of imaging surfaces with atomic resolution for both conducting and nonconducting samples. In the AFM, a very sharp tip is attached to a microcantilever arm. When operating in the contact mode, the tip is brought into contact with the sample surface and it scans along the surface in an x-y raster pattern. Images are constructed pixel by pixel from the height of the tip apex at each x-y position. The AFM used in this work is a Digital Instruments Nanoscope II, using the

^{a)}Electronic mail: mTabet@nanometrics.com

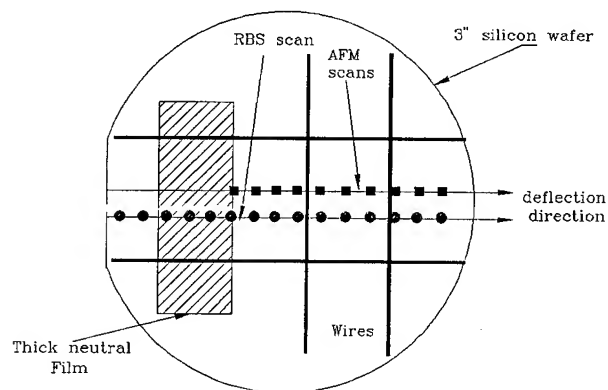


FIG. 1. Schematic of substrate showing the direction of the RBS scan, position of the neutral film, and location of measurement wires.

silicon nitride tips provided with the instrument.

In practice there is distortion in AFM images caused by a nonideal shape and size of probe tips. The size of the tip limits the resolution of the AFM because an infinitely sharp tip is needed to get an accurate picture of a surface with no overhangs. AFM images are a result of the interaction between tip and sample. When feature sizes start approaching the size of the tip, the images obtained contain more distortion, and are said to be a convolution of the tip and sample geometries.^{9,10} When a point other than the apex of the tip touches the sample, distortion is introduced into the image, as shown in Fig. 3. The larger the tip the more distortion the images will contain. This tip effect phenomenon increases the width of scanned features and produces this distortion.¹¹ One way to deal with this problem is to make sharper tips, a second is to remove as much distortion as possible from the images by means of image analysis tools.

Several researchers have presented algorithms for scanning tunneling microscopy (STM)¹²⁻¹⁴ and AFM^{9,10,15-17}

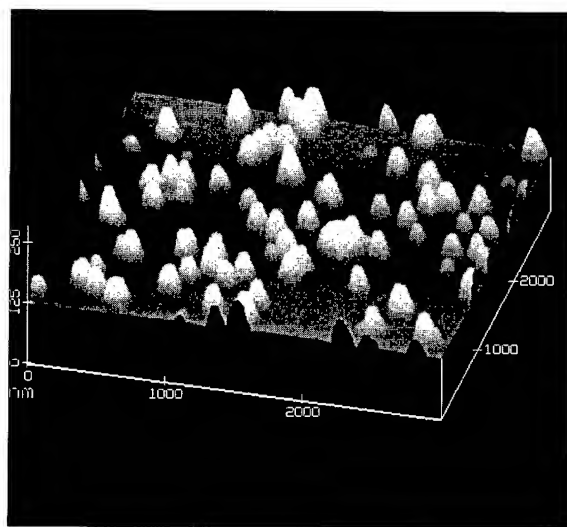


FIG. 2. $3 \times 3 \mu\text{m}$ image of nanometer size zinc islands on a flat silicon substrate. Taken using a nanoscope II (Digital Instruments), running in the constant force mode.

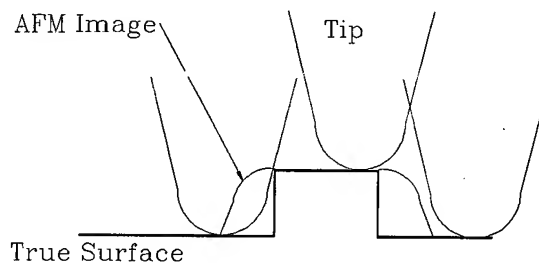


FIG. 3. As the tip scans sharp features it touches the surface at more than one point causing distortion in the resulting image. The width is increased while the height is not.

where an image can be deconvolved using a known tip shape to construct a more accurate picture of the surface. Other methods use a known sample geometry to reconstruct the shape of the probe tip from an AFM image.¹⁸⁻²¹ Also with no knowledge of the scanned surface, feature sizes were used to place an upper bound on the size of the tip.²² A good estimate of the tip shape is obtained if the imaged sample contained sharp features.

II. EXPERIMENT

The AFM images are converted to ASCII data files and imported into the high performance numeric computation software package MATLAB[®] where they consist of two dimensional arrays, in which each pixel value is a number representing the height information in nanometers for that pixel in the image.²³ A program was written in MATLAB[®] to calculate the volume of the islands and the number of zinc atoms present on the surface from the AFM images.²⁴ The number of zinc atoms per unit area was also measured by RBS and compared to AFM results. These two sets of data taken on these nanometer size zinc islands show a qualitative agreement, but there is an offset with the AFM reporting more zinc than RBS. Figure 4 shows the number of atoms/cm² versus deflection distance from RBS and AFM images for a zinc deposition at a crucible temperature of 1260 °C. The peak in the RBS data around 15 mm is due to the neutral zinc beam and is not in the form of islands but is a continuous film. The general shape of the two curves agree, but one possible contribution to the difference is the distortion due to the tip-sample interaction, also called convolution which causes the islands in the images to appear larger. A deconvolution algorithm is required to remove part of this tip effect.

Previously reported convolution and deconvolution algorithms^{9,10} were implemented using MATLAB[®], to study and simulate the interaction between tip and sample, and to remove the distortion (deconvolve) from the zinc islands images, in order to give better agreement between AFM and RBS data. The convolution algorithm simulates the scanning process and produces images with distortion due to tip effects. Then the deconvolution algorithm was developed and tested with these distorted images to examine its effective-

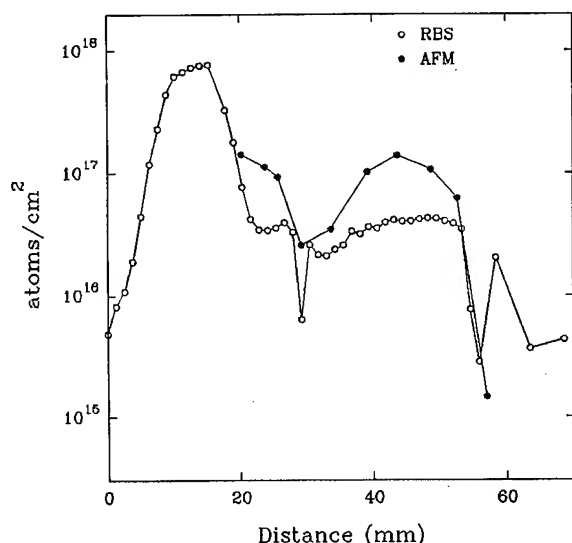


FIG. 4. Atoms per unit area as measured by Rutherford backscattering spectrometry open circles, and AFM, filled circles, for zinc deposited with a crucible temperature of 1260 °C.

ness in reproducing the original surface. The deconvolution algorithm removes part of the distortion resulting in more accurate images.

III. CONVOLUTION AND SIMULATION OF AFM SCANNING

The convolution algorithm uses a simulated tip having a shape similar to the tips provided by Digital Instruments, a pyramid 2 μm high, 4 μm wide at the base, and a 40 nm radius sphere at the end. The process is based on the fact that the AFM constructs images into a two-dimensional array where each cell is the height of the sample which is taken from the position of the tip at that point. Two arrays are used in this procedure, one consisting of the tip, and the other the surface. The pyramid shaped tip is inverted, with the center its lowest point set to zero height, and all other points are greater than zero. The tip is placed on the surface at a corner of the image, and the process starts by raster scanning the surface one pixel at a time under the tip. At each position the difference between the two arrays is taken pixel by pixel, detecting if a point other than the apex of the tip is touching the surface. At each step the tip is moved up or down maintaining the distance between the two surfaces at a minimum. The image is constructed from the position of the tip apex. Simulations were done using several surfaces like tables, steps, spikes, rulings, and spheres. The height of these features was correct but the width increased and the effect of tip-sample convolution was observed.

IV. DECONVOLUTION AND REMOVAL OF DISTORTION FROM IMAGES

The deconvolution algorithm removes part of the distortion introduced by the tip, producing more accurate images. The process starts with two arrays, one containing the tip and

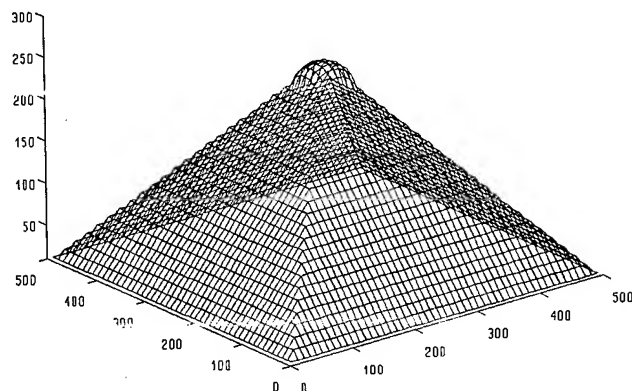


FIG. 5. Mesh of the tip used in the simulations. Units are in pixels.

the other, the scanned and distorted surface. The apex of the simulated tip is placed at each pixel in the image, and areas where the volume of the tip and the surface overlap are removed from the image. This in effect takes into account the physical volume occupied by the tip. This overlapped volume cannot be part of the surface because it was occupied by the tip during scanning. The remaining image is now a more accurate representation of the surface. Areas of the surface not accessed by the tip cannot be recovered by this procedure. These areas result from the tip contacting more than one point simultaneously.

V. RESULTS AND DISCUSSION

Simulation using both the convolution and deconvolution algorithms were carried out using several surface shapes. Figure 5 shows a mesh of a pyramid tip with a sphere at the end, which was used in a simulation of the AFM scanning process. Figure 6(a) is a simulated surface of a table that will be processed by the convolution program. Figure 6(b) is the surface that would be obtained from the AFM if the surface in Fig. 6(a) was scanned by the tip of Fig. 5, all else being ideal. It can be seen from Fig. 6(b) that the surface has been distorted, and the width of the table has increased because of the finite size of the tip, and the sharp edges of the table were rounded by the spherical shape of the tip. When Fig. 6(b) is processed by the deconvolution algorithm some of the distortion is removed from the image leaving a surface, shown in Fig. 6(c), closely resembling the original surface from Fig. 6(a). It is seen from Fig. 6(c) that the original surface could not be reconstructed completely because of areas the tip never scanned.

The images of the nanometer size zinc islands were processed by the deconvolution algorithm; Fig. 7 is Fig. 2 after deconvolution. The number of atoms/cm² was again calculated from these deconvolved images and is shown in Fig. 8 along with the RBS data. Figure 8 also shows a plot where the ZnO layer, which stops growing at a thickness of about 10 nm and does not interfere with RBS measurements, is taken into account. These two plots are in better agreement with the RBS data than the plot in Fig. 4. There are still small disagreements between the data that are due in part to

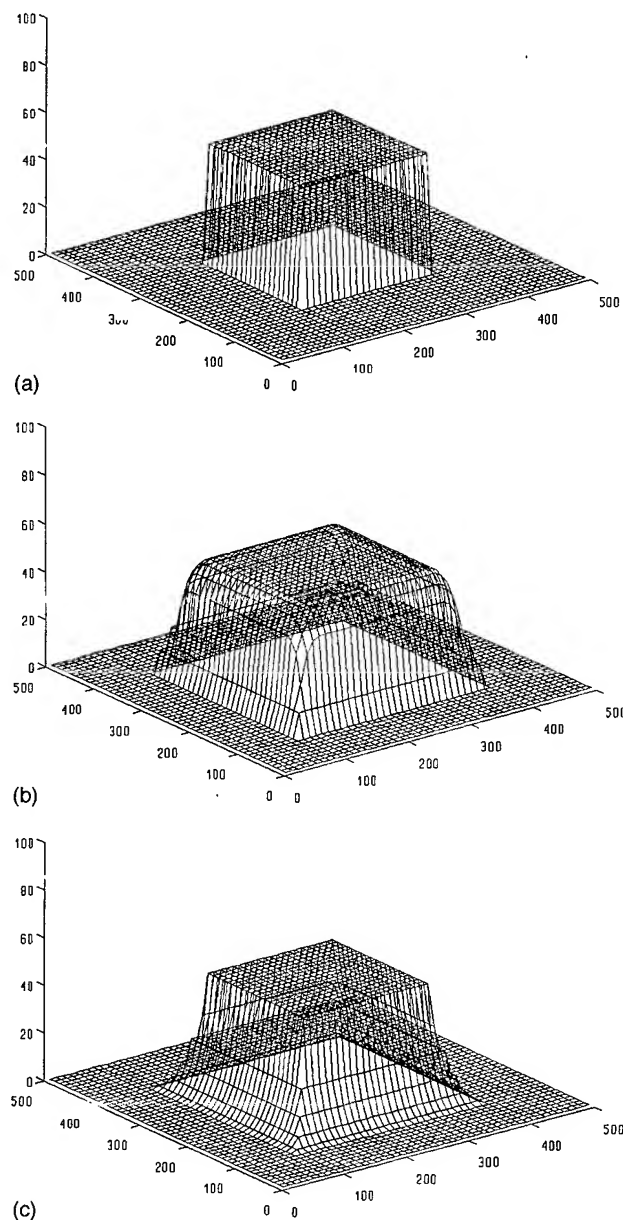


FIG. 6. (a) Simulated surface of a table to be processed by the convolution program. (b) The simulated AFM image that would be obtained if the surface in (a) were scanned by the tip of Fig. 5, all else being ideal. (c) Improved simulated surface when (b) is processed by the deconvolution algorithm, some of the distortion is removed from the image. This closely resembles the original surface from (a).

the approximation made. First, the shape and size of the tip used in scanning the zinc islands images was unknown because the images were scanned several years ago, so a tip resembling the tip described by Digital Instruments was used in the deconvolution procedure. In future uses of the deconvolution program a known sample will be scanned and the tip shape reconstructed from the resulting image. This reconstructed tip would then be used in the deconvolution procedure. Second is the use of bulk zinc atomic density, which could not be confirmed for nanometer size islands, in the calculation. Third is the comparison of data from a $3\ \mu\text{m}$

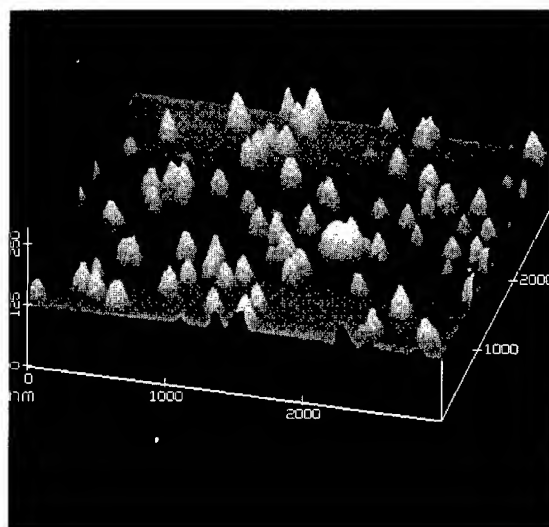


FIG. 7. $3 \times 3\ \mu\text{m}$ image of nanometer size zinc islands on a flat silicon substrate, after processing by the deconvolution algorithm. Taken using a nanoscope II (Digital Instruments), running the constant force mode.

square image and a 1 mm diameter RBS beam spot which could introduce a statistical error.

The deconvolution algorithm was also tested on a Co_3O_4 film deposited by the spray Pyrolysis process for 8 s onto heated 4.8-mm-thick soda-lime float glass. The rms roughness of the film obtained from a $5 \times 5\ \mu\text{m}$ AFM image was 2.62 nm. The rms roughness changed to 2.13 nm after processing the image by the deconvolution program. Since the deconvolution process removes parts of the surface that would be occupied by the tip during the scanning the rms roughness decreased. The appearance of the two surfaces

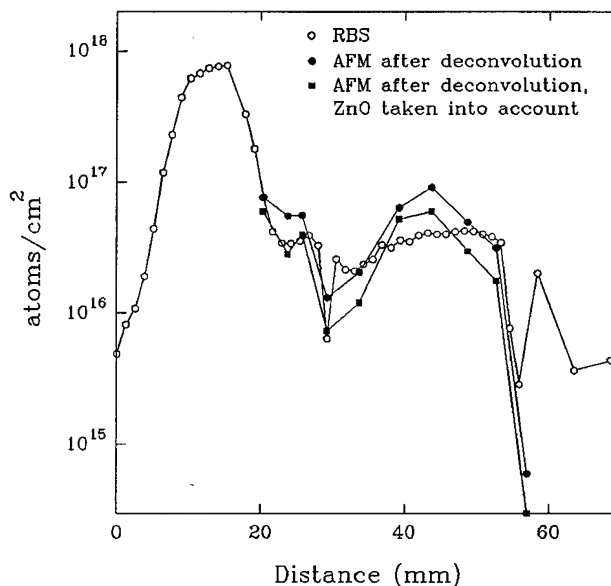


FIG. 8. Atoms per unit area as measured by RBS (open circles) and deconvolved AFM images (filled circles) for zinc deposited with a crucible temperature of $1260\ ^\circ\text{C}$. Filled squares indicate that the ZnO layer was taken into account in the AFM images.

were similar and the change in rms roughness was not significant indicating that the tip was sharper than most surface features.

VI. CONCLUSIONS

A comparison of AFM and RBS data taken on nanometer size zinc islands on a flat substrate agree qualitatively but the AFM reports more zinc than RBS. The distortion caused by the nonideal shape and size of AFM probe tips was removed from the images of nanometer size zinc islands by means of a deconvolution algorithm. The comparison made after deconvolution of the AFM images showed better agreement between the data confirming the usefulness of algorithms.

Presented at the 43rd National Symposium of the American Vacuum Society, Philadelphia, PA, 12–18 October 1996.

¹T. Takagi *Ionized Cluster Beam Deposition and Epitaxy* (Noyes, Park Ridge, NJ, 1988).

²I. Yamada, H. Usui, and T. Takagi, *J. Phys. Chem.* **91**, 2463 (1987).

³W. L. Brown, M. F. Jarrold, and H. Usui, *Nucl. Instrum. Methods Phys. Res. B* **59**, 182 (1991).

⁴F. K. Urban III, *J. Appl. Phys.* **67**, 7082 (1990).

⁵F. K. Urban III and A. Bernstein, *Thin Solid Films* **193/194**, 92 (1990).

⁶F. K. Urban III, S. W. Feng, and J. J. Nainaparampil, *Appl. Phys. Lett.* **61**, 1190 (1992).

⁷F. K. Urban III, S. W. Feng, and J. J. Nainaparampil, *J. Vac. Sci. Technol. B* **11**, 1916 (1993).

⁸F. K. Urban III, S. W. Feng, and J. J. Nainaparampil, *J. Appl. Phys.* **74**, 1335 (1993).

⁹P. Markiewicz and M. C. Goh, *J. Vac. Sci. Technol. B* **13**, 1115 (1995).

¹⁰P. Markiewicz and M. C. Goh, *Langmuir* **10**, 5 (1994).

¹¹U. Schwartz, H. Haefke, P. Reimann, and H. Gunthrod, *J. Microsc.* **173**, 183 (1993).

¹²R. Chicon, M. Ortuno, and J. Abellan, *Surf. Sci.* **181**, 107 (1987).

¹³P. Niedermann and O. Fischer, *J. Microsc.* **152**, 93 (1988).

¹⁴G. Reiss, F. Schneider, J. Vancea, and H. Hoffmann, *Appl. Phys. Lett.* **57**, 867 (1990).

¹⁵P. Markiewicz and M. C. Goh, *Rev. Sci. Instrum.* **66**, 3186 (1995).

¹⁶D. Keller and F. Franke, *Surf. Sci.* **268**, 333 (1992).

¹⁷D. Keller and F. Franke, *Surf. Sci.* **294**, 409 (1993).

¹⁸R. Miller, J. Vasenska, and E. Henderson, *SIAM (Soc. Ind. Appl. Math.) J. Appl. Math.* **55**, 1362 (1995).

¹⁹P. Godowski, V. Maurice, and P. Marcus, *Chem. Anal.* **40**, 231 (1995).

²⁰J. Vasenska, R. Miller, and E. Henderson, *Rev. Sci. Instrum.* **65**, 2249 (1994).

²¹L. Montelius and J. Tegenfeldt, *Appl. Phys. Lett.* **62**, 2628 (1993).

²²J. S. Villarrubia, *Surf. Sci.* **321**, 287 (1994).

²³C. S. Lindquist and F. K. Urban III, *Thin Solid Films* **270**, 399 (1995).

²⁴M. F. Tabet and F. K. Urban III, *Thin Solid Films* **290–291**, 312 (1996).

Effect of substrate temperature and annealing on the structural properties of ZnO ultrafine particle films

Zhao Dachun^{a)}

Department of Physics, East China University of Science and Technology, Shanghai 200237, People's Republic of China

Qu Zhongkai

Shanghai Petrochemical Institute, Shanghai 200540, People's Republic of China

Pan Xiaoren

Department of Physics, East China University of Science and Technology, Shanghai 200237, People's Republic of China

Dai Muji and Sun Minggen

Analysis and Research Centre, East China University of Science and Technology, Shanghai 200237, People's Republic of China

(Received 25 June 1996; accepted 4 April 1997)

ZnO ultrafine particle (UFP) films were deposited on glass and SiO₂ substrates by a direct current gas discharge activated reactive evaporation method. The effect of substrate temperature and annealing on the structure and morphology of ZnO UFP films was studied by x-ray diffraction and scanning electron microscope. The results show that the spherical island density decreases with increasing annealing temperature and the structure becomes polycrystalline with a (002) preferential orientation as the substrate temperature increases. In addition, angle resolved x-ray photoelectron spectroscopy was used to study the absorption of water on the ZnO UFP film surface by measuring the two deconvoluted peaks for O 1s. The two deconvoluted peaks for O 1s were located at 533.2 and 534.8 eV. The absorption coefficients of water on the surface were 0.52 and 0.43, respectively, for nonannealed and annealed ZnO UFP films. © 1997 American Vacuum Society. [S0734-211X(97)00204-7]

I. INTRODUCTION

In recent years, metal and metal oxide ultrafine particle (UFP) films have received greater attention as catalysts, gas sensors, and gas separation.^{1,2} Significant effort has been made to understand the surface morphology and structure, preparation, as well as electrical and optical properties of ultrafine particle films under various conditions.³⁻⁵ ZnO has a wide gap near 3.3 eV and is an *N*-type semiconductor that crystallizes in a hexagonal wurtzite structure with lattice parameters $a = 3.25$ and $c = 5.21$ Å.^{1,3} It has many potential applications in energy efficient windows, liquid crystal displays, and as a transparent electrode material. At present, although many techniques have been developed to prepare ZnO UFP films, it is difficult to obtain highly purity, transparent, conducting ZnO UFP films. In this article, the structure and morphology of ZnO UFP films deposited on glass and SiO₂ substrates at different substrate temperatures and annealing conditions are reported. The absorption of water on the ZnO UFP film surface was also investigated by angle resolved x-ray photoelectron spectroscopy (ARXPS).

II. EXPERIMENTAL DETAILS

ZnO UFP films were deposited on glass and SiO₂ substrates using a direct current (dc) gas discharge activated re-

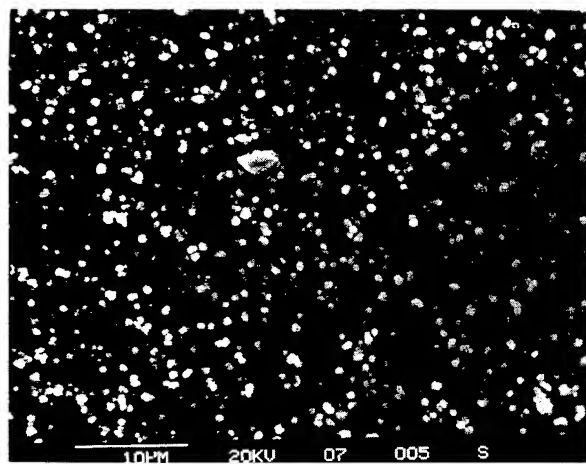
active evaporation (ARE) method. The preparation conditions and details of the dc gas discharge ARE deposition systems are described elsewhere.^{3,6} Zinc (99.9%) was used as the source material. The ZnO UFP films were prepared under an oxygen gas discharge forming plasma. The morphology and structure of ZnO UFP films were studied by scanning electron microscopy (SEM) and x-ray diffraction (XRD). X-ray diffraction measurements were carried out on a Rigaku diffractometer using a Cu $K\alpha$ ($\lambda = 1.54$ Å) source line (operating at 40 kV and 40 mA) and a scanning speed of 4° per min. Angle resolved x-ray photoelectron spectroscopy (ARXPS) with a Mg $K\alpha$ (1253.6 eV) x-ray source operating at 250 W power was used to study the physical absorption of water on the ZnO UFP film surface. The spectrometer pass energies were 89.45 and 35.75 eV for the survey spectra and the detailed region, respectively.

III. RESULTS AND DISCUSSION

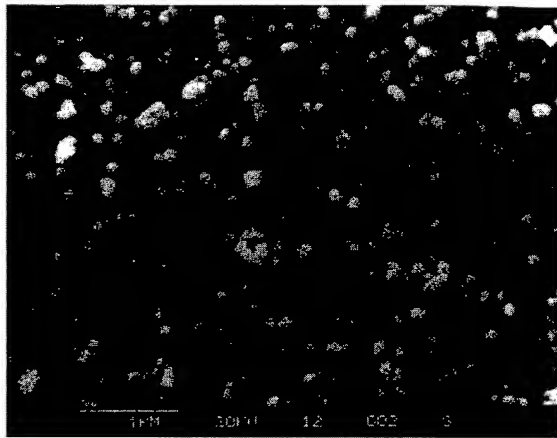
A. Influence of conditions on the morphology and structure

Substrate temperature and annealing have a strong influence on the morphology and structure of ZnO UFP films. Figures 1(a)–1(c) show SEM photographs of ZnO UFP films deposited on glass at different substrate temperatures (room temperature and 250 and 350 °C). As can be seen, small nonuniform spherical islands are distributed across the smooth glass substrate, and the spherical islands aggregate

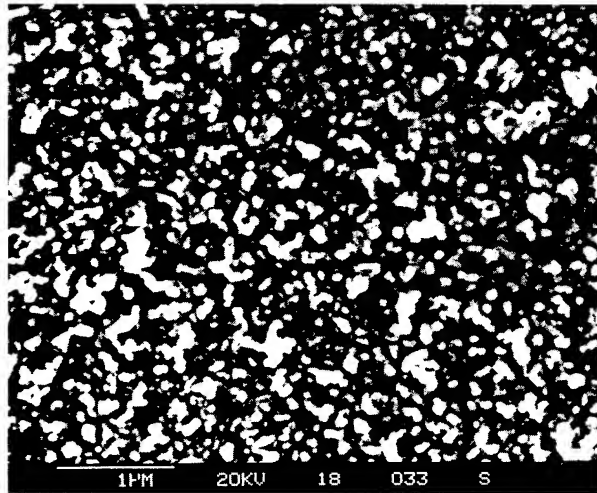
^{a)}Present address: 720 W 30th St. #3B, Los Angeles, CA 90007; Electronic mail: lxu@scf.usc.edu



(a)



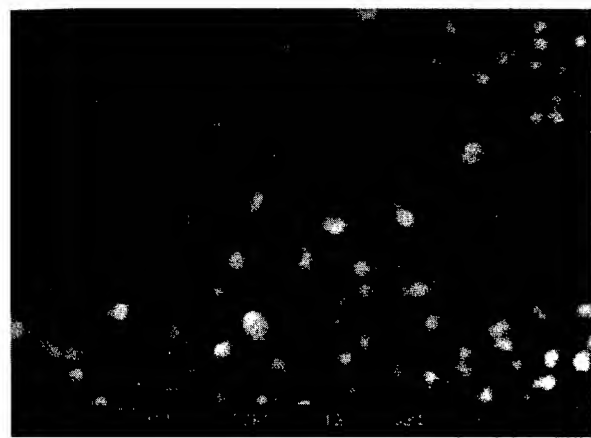
(b)



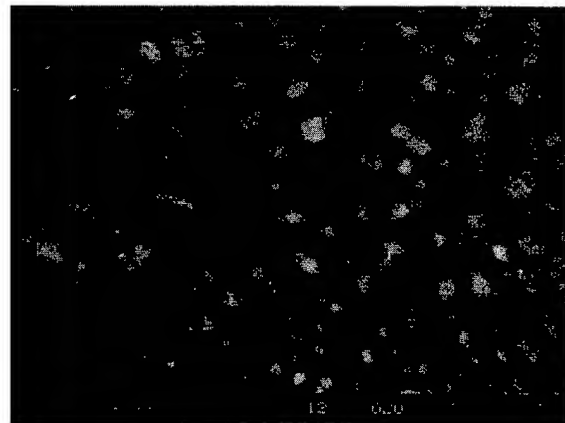
(c)

FIG. 1. The morphology of ZnO UFP films at different temperatures. (a) Room temperature, (b) 250 °C, (c) 350 °C.

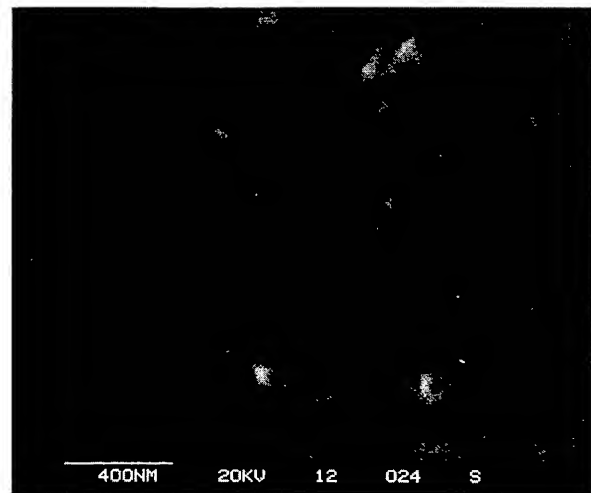
with increasing substrate temperature. Figures 2(b) and 2(c) show SEM images of ZnO UFP films deposited on a SiO₂ substrate at annealing temperatures of 350 and 550 °C, respectively. It can be seen that the spherical island density decreases as the annealing temperature increases, and the



(a)



(b)



(c)

FIG. 2. The morphology of nonannealed and annealed ZnO UFP films. (a) Not annealed ZnO UFP films and (b) and (c) annealed at 350, and 550 °C for 2 h in air.

island size increases after annealing at 550 °C for 2 h. The reason for this is that the spherical island can obtain enough energy to move and to aggregate on the substrate surface during annealing.

X-ray diffraction patterns of ZnO UFP films at various

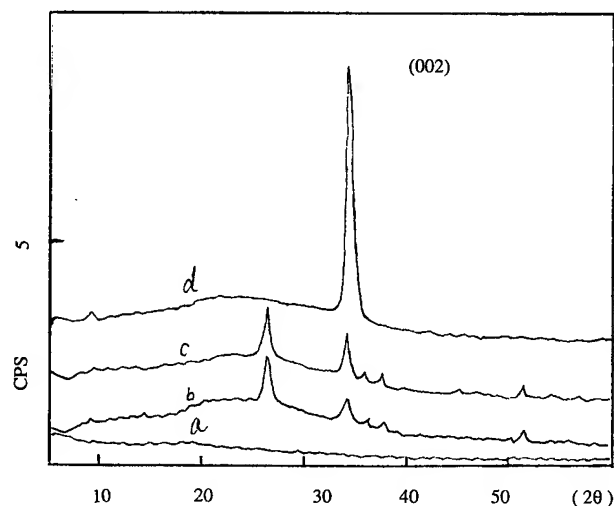


FIG. 3. XRD spectrum for ZnO UFP films at different substrate temperatures. (a) Room temperature and (b) 160, (c) 250, (d) 350 °C.

substrate temperatures (room temperature and 160, 250, and 350 °C) are shown in Figs. 3(a)–3(d). The results indicate that the ZnO UFP films exhibit an amorphous structure at low substrate temperature; however, the ZnO UFP films exhibit a polycrystalline structure with a (002) preferential orientation with increasing substrate temperature. The results can be confirmed by calculating the diffraction intensity of ZnO UFP films at different substrate temperatures. The diffraction intensity ratio is $I_1(002):I_2(002):I_3(002)=61:78:100$ at substrate temperatures of 160, 250, and 350 °C, respectively. In addition, the linewidth ratio $W1(002):W2(002):W3(002)=48:45:54$, and the lattice spacing ratio $d1(002):d2(002):d3(002)=2.609:2.616:2.607$ are obtained from measured data. Figures 4(a)–4(d) show the structure of the annealed UFP films change from amorphous to polycrystalline. The peaks at $2\theta = 31.7, 34.3, 36.2$, and 47.5 , that correspond to (100), (002), (101), and (102), respectively, are shown in Fig. 4(d). The structure of ZnO UFP films tends towards polycrystalline in the wurtzite structure at a high annealing temperature of 500 °C.

B. ARXPS and physical absorption

Angle resolved x-ray photoelectron spectroscopy is a useful tool to examine the absorption state. Table I summarizes

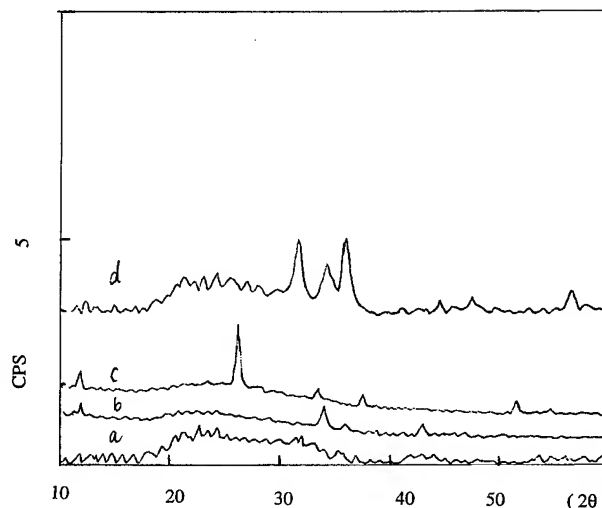


FIG. 4. Spectra of ZnO UFP films. (a) Not annealed ZnO UFP films and (b), (c), and (d) annealed ZnO UFP films at 200, 300, and 500 °C.

the peak positions and atomic concentrations of Zn 2P 3/2, O 1s and C 1s lines of the ZnO UFP films before and after annealing. One can see from Table I that the oxygen to zinc ratio decreases as the polar degree increases, and the oxygen to zinc ratio is derived from chemical stoichiometry. This can be explained by the fact that there are many oxygen vacancies near the surface of ZnO UFP films, and there are many zinc interstitials inside the surface of ZnO UFP films. Figures 5(a) and 5(b) show detailed O 1s ARXPS of ZnO UFP films before and after annealing. It is obvious from Fig. 5(a) that two deconvoluted peaks of O 1s were found, and the positions of the two deconvoluted peaks of O 1s are located at 533.2 eV (O_I), and 534.8 eV (O_{II}) binding energies, respectively. The lower binding energy at 533.2 eV is attributed to the ZnO UFP film's O 1s; the higher binding energy at 534.8 eV results from the absorption of water on the ZnO UFP film surface. For UFP films, the film porosity (P) can be expressed as follows:⁷

$$P = \Theta \frac{O_{II}}{O_I + O_{II}}, \quad (1)$$

where Θ is absorption coefficient, and O_I and O_{II} are the atomic concentrations of the two deconvoluted peaks of O 1s. The relation between the optical parameters and film packing density can be expressed using⁸

TABLE I. The ARXPS data for the ZnO UFP films (A: before annealing; B: after annealing at 500 °C for 2 h; BE: binding energy; C: atomic concentration).

A							B				
		Zn2p $\frac{3}{2}$	O _I 1s	O _{II} 1s	C1s	O _I /Zn	Zn2p $\frac{3}{2}$	O _I 1s	O _{II} 1s	C1s	O _I /Zn
10°	BE	1021.7	533.2	534.8	284.6	0.83	1021.9	533.0	535.1	284.7	0.85
	C	(37.0%)	(30.8%)	(26.5%)	(20.0%)		(31.2%)	(26.6%)	(34.5%)	(17.0%)	
45°	BE	1021.9	531.6	534.2	284.6	0.79	1021.9	531.6	534.2	284.6	0.79
	C	(48.0%)	(37.5%)	(17.7%)	(15.6%)		(45.7%)	(35.4%)	(10.4%)	(15.0%)	
	BE	1022.1	531.4	534.2	284.6	0.74	1022.1	531.4	534.2	284.6	0.74
	C	(51.1%)	(38.4%)	(10.8%)	(9.7%)		(47.9%)	(36.0%)	(8.8%)	(8.2%)	

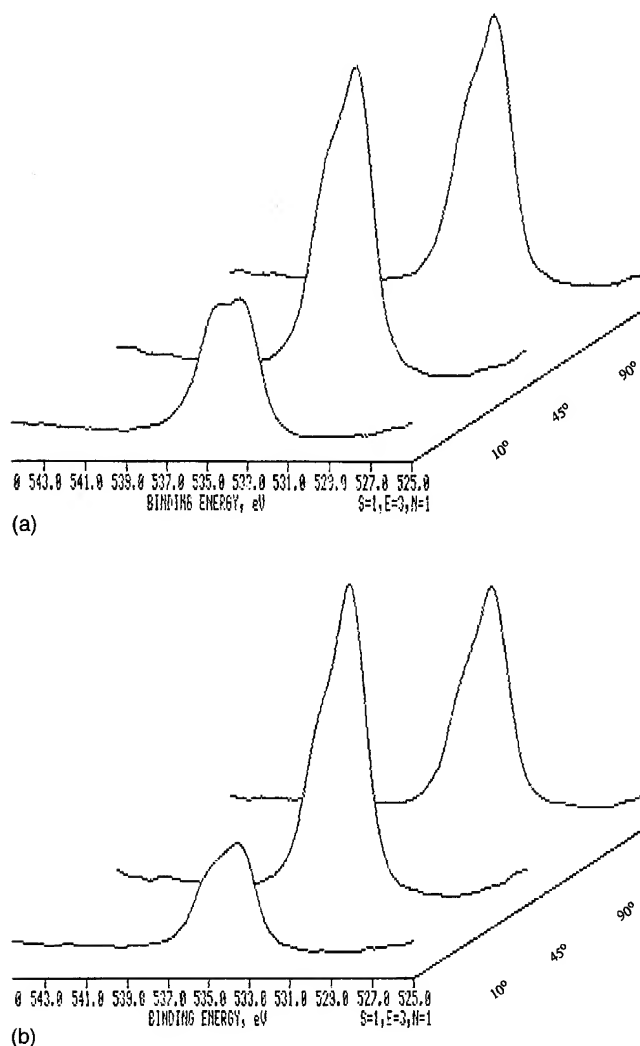


FIG. 5. Detailed O 1s ARXPS of ZnO UFP films before and after annealing. (a) Before annealing and (b) after annealing at 300 °C for 2 h.

$$\frac{n_f^2 - 1}{n_f^2 + 2} = \frac{n_b^2 - 1}{n_b^2 + 2} P_d + \frac{n_w^2 - 1}{n_w^2 + 2} a(1 - P_d); \quad (2)$$

here, n_f , n_b , and n_w are the refraction indices of the film, bulk materials, and water, respectively. For Eq. (2) a is a coefficient that indicates the relative proportion of the pores filled with water, and was taken as 1, meaning all the pores were filled with water. For UFP films, $P_d = 1 - P$, n_b and n_w are 2.013 and 1.33, respectively.⁹ The following equations were used to calculate $n_f(\lambda)$ from infrared (IR) spectra transmittance characteristics:

$$n_f(\lambda) = \frac{1}{2} \{ 8n_0(\lambda)c(\lambda) + [n_0(\lambda) + 1]^2 \}^{1/2} + \frac{1}{2} \{ 8n_0(\lambda)c(\lambda) + [n_0(\lambda) - 1]^2 \}^{1/2}, \quad (3)$$

$$n_0(\lambda) = \frac{1 + [1 - T_0(\lambda)]^{1/2}}{T_0(\lambda)}, \quad (4)$$

$$C(\lambda) = \frac{T_{\max}(\lambda) - T_{\min}(\lambda)}{2T_{\max}(\lambda)T_{\min}(\lambda)}, \quad (5)$$

where $n_0(\lambda)$ is the refractive index of the substrate, $T_{\max}(\lambda)$ and $T_{\min}(\lambda)$ are experimentally traced envelope curves passing on the maxima and minima. The values of $T_0(\lambda)$ and $T_{\max}(\lambda)$ are close in the very weak absorption region. From the IR spectra, n_f of 1.8114 ($\lambda = 500$ nm wavelength) for ZnO UFP films (particle size 80 nm) is obtained using Eqs. (3), (4), (5) and our previous IR results for ZnO UFP films. Therefore, the absorption coefficient of water (Θ) on the ZnO UFP film surface ($\theta = 10^\circ$) at different conditions can be calculated by Eqs. (1) and (2). The Θ values are 0.52 and 0.43 for nonannealed and annealed ZnO UFP films, respectively. The results show that the absorption coefficient of water on the ZnO UFP film surface decreases after annealing due to the decreasing specific surface ratio.

IV. CONCLUSIONS

ZnO UFP films deposited on glass and SiO₂ substrates by a dc gas discharge activated reaction were characterized by SEM, XRD, and ARXPS. The effect of substrate temperature and annealing on the structure and morphology of ZnO UFP films was described. The results show the spherical island density decreases as annealing temperature increases. The diffraction intensity ratio is $I_1(002):I_2(002):I_3(002) = 61:78:100$ at different substrate temperatures, and it exhibits a structure with (002) preferential orientation with increasing substrate temperature. The islands aggregate during annealing and tend to transform from amorphous to crystalline. ARXPS analysis of the ZnO UFP film was also used to study the absorption of water on the ZnO UFP film surface.

¹R. Lalauze, P. Breuil, and C. Pijolat, *Sens. Actuators B* **3**, 175 (1991).

²G. Camisio, F. Gerbassi, G. Petrini, and G. Parravano, *J. Catal.* **54**, 66 (1978).

³D. Zhao and X. Pan, *J. Inst. Chem. Technol.* **3**, 383 (1993).

⁴L. G. Mar, P. Y. Timbrell and R. N. Lamb, *Thin Solid Films* **223**, 341 (1993).

⁵D. Zhao and X. Pan, *J. Vac. Sci. Technol. B* **12**, 2880 (1994).

⁶Y. Zhu, H. Lu, Y. Lu, and X. Pan, *Thin Solid Films* **224**, 82 (1993).

⁷L. Meng, C. P. Moreira deSa, and M. P. dos Santos, *Appl. Surf. Sci.* **78**, 557 (1994).

⁸G. Atanassov, R. Thielsch, and D. Popov, *Thin Solid Films* **223**, 288 (1993).

⁹*Handbook of Chemical Physics*, 57th ed., edited by R. C. Weast (Chemical Rubber, Cleveland, OH, 1976).

Ion implanted nanostructures on Ge(111) surfaces observed by atomic force microscopy

Y. J. Chen,^{a)} I. H. Wilson, W. Y. Cheung, J. B. Xu, and S. P. Wong

Department of Electronic Engineering and Materials Technology Research Centre, The Chinese University of Hong Kong, Shatin, N. T., Hong Kong

(Received 20 November 1996; accepted 11 April 1997)

Epi-ready Ge(111) surfaces were implanted with cobalt ions to doses of 10^{16} – 5×10^{17} ions/cm² at accelerating voltages of 40–70 kV. Cellular nanostructures were observed by contact mode and tapping mode atomic force microscopy (AFM). These are similar (at higher resolution) to those reported in earlier scanning electron microscope measurements. Image distortions observed in contact mode AFM are attributed to not only the effect of the tip size but also the change of the effective tip shape due to the softness and stickiness of the implanted surface layer. The variation of the root-mean-square roughness with ion dose (10^{16} – 10^{17} ions/cm²), accelerating voltage (40–70 kV), and mean beam current density (15–150 μ A/cm²) is presented and explained in terms of ion range and surface temperature. © 1997 American Vacuum Society. [S0734-211X(97)03404-5]

I. INTRODUCTION

Germanium surfaces bombarded to high ion doses have been observed by scanning electron microscopy (SEM)¹ and transmission electron microscopy (TEM).^{2–5} Both techniques showed that a complex cellular or sponge-like swollen structure had evolved from surface holes or cavities with increasing ion dose. However, not much has been done further to study this abnormal radiation-induced swelling of amorphous materials and the mechanisms leading to the formation of this structure are not well understood. Recently, we used atomic force microscopy (AFM) to observe these unique damage structures on self-ion bombarded germanium samples that one of the authors had kept for fourteen years since earlier SEM measurements. Not only cellular structures but also more generalized columnar structures were found by AFM; these structures were further confirmed by observations of the surface of a newly implanted Ge(111) sample (5×10^{16} Ge⁺/cm² at 100 keV).⁶

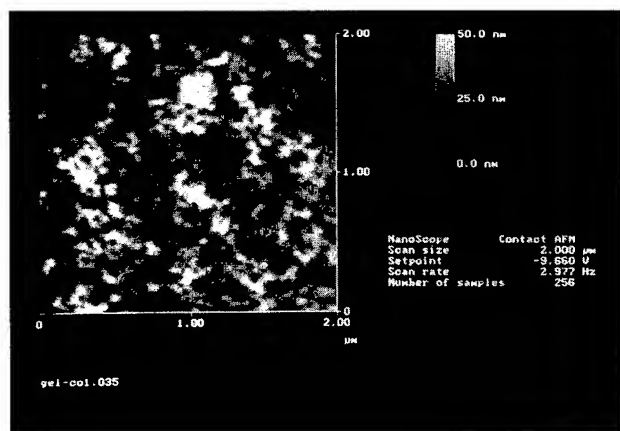
In this article, we present an investigation of cobalt (Co) implanted Ge(111) surfaces by AFM. It is well known that AFM can image conductive and nonconductive surfaces with higher resolution than SEM and TEM with nearly no sample preparation. Bombarded nanostructures on germanium surfaces are an appropriate size for AFM. More important, the AFM is a profilometer-based technique; the width and height measurements in the images are quantitative. These quantitative data allow simple, accurate and quantitative characterization of surfaces and surface structures. Among the quantitative parameters, surface roughness is important for indicating the dynamic response of the surface to the bombardment. In this study, special attention is therefore given to the investigation of the relationship of root-mean-square roughness (R_q) to ion energy, ion dose, beam current density and temperature. Roughness gives complementary information on the formation mechanisms of these nanostructures. AFM images are subject to some artifacts of their own

mainly due to the tip size, however. Severe image distortions were observed when the relative sharpness of the tip was comparable to that of surface features. The undesirable artifacts have to be identified and avoided in the measurements. With the most newly developed tapping mode operation, the image artifacts presented in contact mode AFM observation of the nanostructures were minimized. Cobalt was chosen as the implant species because of its ability to form germanic nanometer magnetic materials (GeCo) by the ion beam synthesis method. Also, the mass of cobalt is high enough to create a collision cascade dense enough to generate structures similar to those observed for self-ion bombardment.

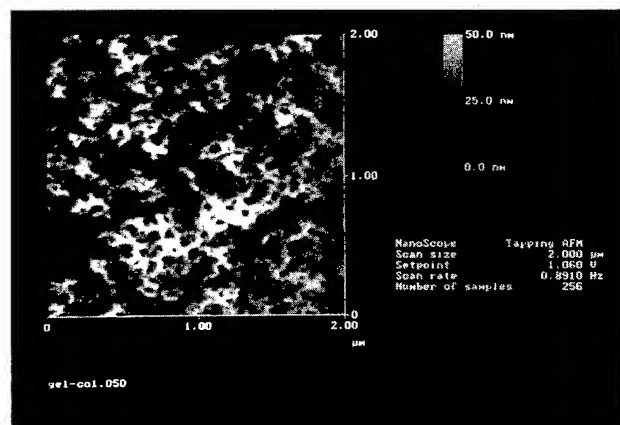
II. EXPERIMENT

The specimens used in this study were small pieces cut from epi-ready (111) 254 μ m-thick wafers of *p*-type (1.4–2.2 Ω cm) germanium (Atramet Inc.). (Some Ge(100) surfaces were also used for implantation for comparison in some cases.) The as-supplied surfaces were generally very smooth and the roughness (R_q) over an area of 1 μ m \times 1 μ m is measured by AFM to be below 0.1 nm. This is trivial compared with the roughness of implanted samples. The Co implantation was carried out in a JYZ-8010W metal vapour vacuum arc (MEVVA) ion implanter⁷ at The Chinese University of Hong Kong. Our small samples were put into the center of the implantation area in order to obtain uniform exposure. The bombardment was conducted at an accelerating voltage of 40 kV–70 kV with a mean beam current of 0.5–7 mA, corresponding to a current density of 15–210 μ A/cm² by taking the implantation area to be ~ 30 cm² (which means a rather high flux in the range of 10^{14} – 1.5×10^{15} cm⁻² s⁻¹). The nominal implantation dose ranged from 10^{16} to 5×10^{17} ions/cm² by measuring the charge collected by the target. (Because of the three charge/energy states for Co ions in the MEVVA source, the nominal dose is the weighted sum of the doses of Co⁺, Co⁺⁺, and Co⁺⁺⁺ ions taking the abundance of the measured species into account.) The change of temperature due to beam heating was

^{a)}Electronic mail: yichen@ee.cuhk.edu.hk



(a)



(b)

FIG. 1. A contact mode (a) and a tapping mode (b) AFM image of a Co implanted Ge(111) sample (dose: 5×10^{16} ion/cm²; accelerating voltage: 70 kV; beam current density: $15 \mu\text{A}/\text{cm}^2$).

monitored by a thermocouple attached to the target holder. The vacuum during implantation was typically better than 10^{-6} Torr.

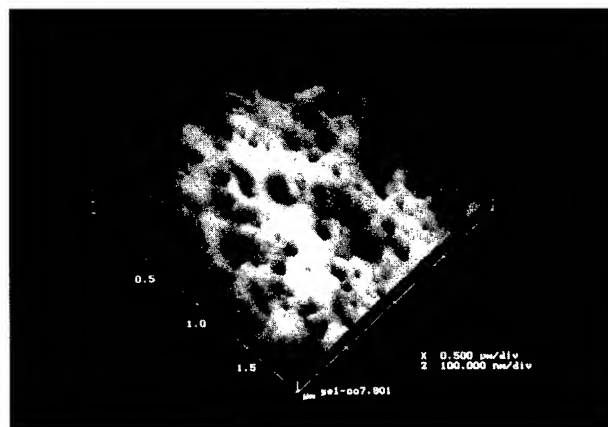
The AFM observations were made in air at room temperature using a standard atomic force microscope and an extended multimode scanning probe microscope (Digital Instruments Nanoscope III, Santa Barbara, California). The former was operated in constant contact force mode with a large field $125 \mu\text{m} \times 125 \mu\text{m}$ piezo head and a standard silicon-nitride (Si_3N_4) cantilever while the latter operated in tapping mode with the same piezo head and a Nano-probeTM (TESP) silicon cantilever. Only offline image processing of first order flattening and third order plane fit was performed throughout using Nanoscope III software.

III. RESULT AND DISCUSSION

Figure 1 shows a contact mode and a tapping mode AFM image ($2 \mu\text{m} \times 2 \mu\text{m}$) of a Co implanted Ge(111) sample (dose: 5×10^{16} ions/cm²; accelerating voltage: 70 kV; beam current density: $15 \mu\text{A}/\text{cm}^2$). Cellular structures appear in both images, in good agreement with the SEM/TEM results in the literature.¹⁻⁵ To corroborate the results, we also used a



(a)



(b)

FIG. 2. A contact mode (a) and a tapping mode (b) three dimensional AFM image of another Co implanted Ge(111) sample (dose: 10^{17} ions/cm²; accelerating voltage: 70 kV; beam current density: $15 \mu\text{A}/\text{cm}^2$).

commercial scanning electron microscope (Amary) to observe the structure and obtained similar results at lower resolution. From Fig. 1, the mean cell diameter is measured to be about 50 nm which is a little smaller than that of a 50 keV Ge^+ implanted surface with the same dose.¹ Cellular structures were observed for other Ge(111) samples and for Ge(100) samples as well.

Figure 2 shows three dimensional AFM images of another sample implanted with a higher dose (10^{17} ions/cm²; 70 kV; $15 \mu\text{A}/\text{cm}^2$). There are some differences between the contact mode and tapping mode AFM images shown in Fig. 2(a) and Fig. 2(b). For tapping mode, the topography [Fig. 2(b)] shows large cells with finer dendritical walls. The mean cell diameter is about 110 nm. For contact mode, the morphology [Fig. 2(a)] is more like a columnar structure with cell walls consisting of overlapped grains or clusters. Careful investigation shows that the contact mode image is more distorted from the real morphology. Reasons for this may be due to (1) larger tip size and lower aspect ratio of the silicon nitride probe compared with the silicon probe; (2) softness of the implanted surface layer; (3) friction between the tip and the surface. To identify and avoid the image artifacts, we will discuss each of these factors.

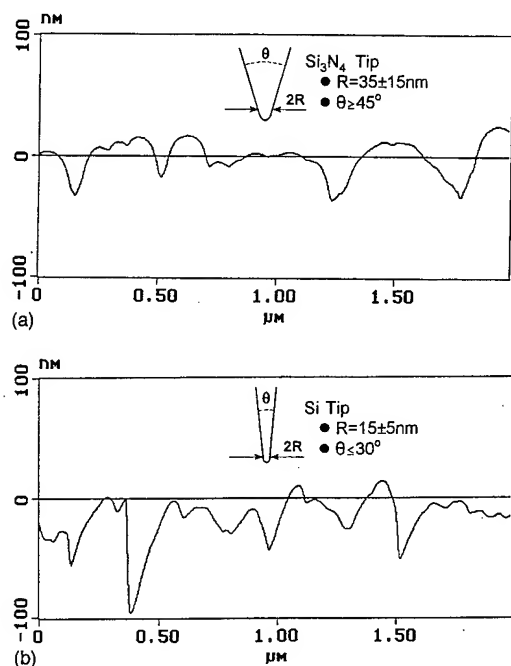


FIG. 3. Single-line scans from Figs. 2(a) and 2(b). The shapes of the silicon nitride tip and the silicon tip are schematically shown in (a) and (b), respectively. The as-shown tips look elongated due to different horizontal and vertical scales.

(1) It is well known that scanning probe microscopy (SPM) is a probe-related technology and the image is the convolution of the tip shape and the true sample morphology.⁸ When the relative sharpness of the tip is comparable to that of surface features, the image can be severely distorted. Observations of image distortions due to the finite size of the AFM tip have been reported for nanometer structures.⁹⁻¹³ Figure 3 shows two single-line scans from Figs. 2(a) and 2(b). The nominal radius of curvature of the silicon nitride tip is 35 ± 15 nm while that of silicon tip is 15 ± 5 nm. The shape of the tip used in contact mode and tapping mode AFM is shown schematically in Fig. 3. It can be seen that the sharper silicon tip in tapping mode AFM can penetrate into the cells more deeply and therefore can produce a less-distorted topographic image.

(2) It should be mentioned that we have observed morphology changes during scanning in contact mode AFM observations. As shown in Fig. 4, with the Si_3N_4 tip scanning from front to back (from bottom to top in the image), the granular or columnar structure (the same as in the last images) in the lower part of the image changed to a cellular structure in the rest of the image (the same as in subsequent images and as in Fig. 1). This was double checked using another tip to ensure that it was not due to poor adjustment of the scan parameters (scan rate, feedback gain, etc.). Since the cellular structure appeared from the columnar structure sometimes only after a few scans, there is a possibility that the tip "swept" the scan area clear and then uncovered the true morphology. This can be ruled out because, once the cellular structure appears, a similar structure was immediately observed at another location. Also, the cellular struc-

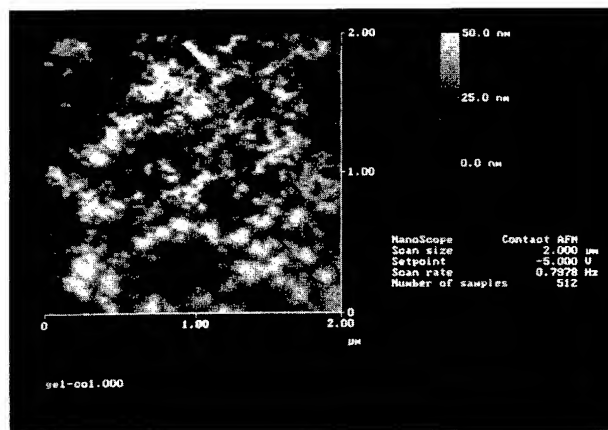


FIG. 4. The apparent morphology changes from the bottom to the top of the image (for the same sample as in Fig. 1).

ture could be imaged on the first pass in some cases. On the other hand, we noted that the cellular structure could change into the columnar structure when the contact force between the tip and the surface was increased above a critical value by raising the feedback setpoint. This cannot be explained by the effect of finite tip size. It seems that either the tip has stuck to material from the surface to form a new tip during scanning or that the shape of the contact area between the tip and surface has been changed/deformed by the loading force. To understand this, we have to investigate the physical properties of the implanted surface. The measurements of elastic-shear-modulus changes by Brillouin scattering¹⁴ indicate a softening of the implanted surface layers. Rutherford backscattering (RBS)/ion channelling analyses,^{2,3} also supported by our x-ray photoelectron spectroscopy (XPS) results,¹⁵ have shown a drastic increase of surface area which is responsible for a higher carbon and oxygen concentration in the near surface layer when exposed to the air, and therefore a blackening of the surface. This may cause an increase in the adhesion or stickiness of the surface. For tapping mode AFM, the interaction between the tip and the surface is greatly reduced (a fraction of a nN) compared with contact mode AFM and therefore produces a less-distorted morphological image. We suggest that the effective tip shape changes due to the softness and stickiness account for the change of the contact mode AFM images. The contact mode therefore yields extra information on the changes in surface properties such as adhesion and elasticity induced by ion bombardment.

(3) The effect of friction should be minor because clear images were obtained after putting the forward and backward traces in coincidence operating in the dual-trace scope mode.

It is interesting that cellular structures have never been observed for implanted silicon samples under comparable conditions.⁷ The unique damage structure in the case of germanium can be understood from two points of view. From sectional SEM/TEM observation, the swollen surface layers consist of elongated holes and honeycombed columns. From the plan-view AFM/SEM/TEM observations, the surface morphology shows a complex cellular or sponge-like, or po-

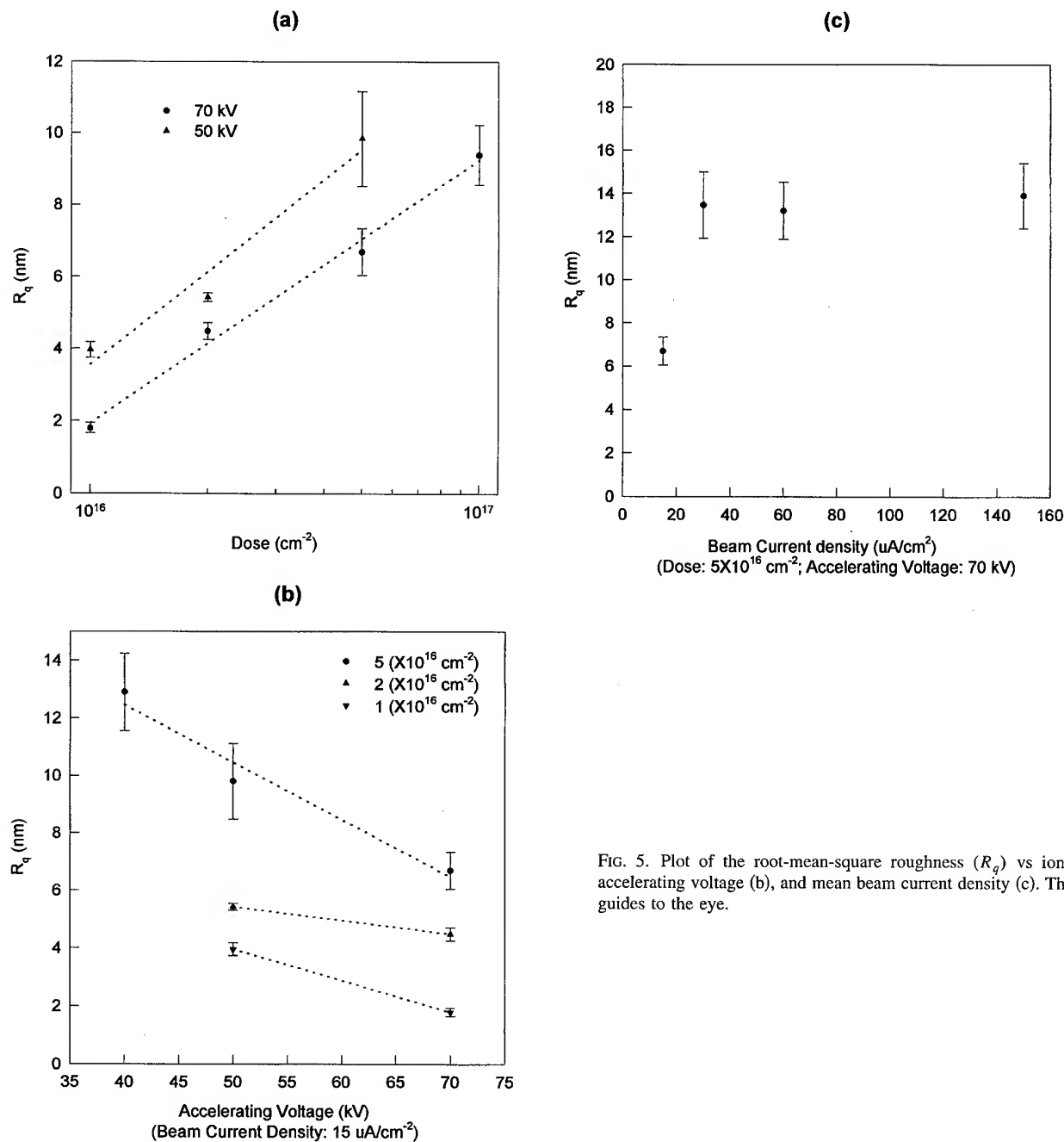


FIG. 5. Plot of the root-mean-square roughness (R_q) vs. ion dose (a), accelerating voltage (b), and mean beam current density (c). The lines are guides to the eye.

rous, structure. (We will show elsewhere in detail that the granular or columnar structure observed by AFM may just be image distortion from the cellular structure.) The softness and large surface area of this amorphous layer would have a deleterious effect on electronic devices but they may have some possible technological applications (such as where a large surface area/volume ratio is needed).

Two mechanisms have been proposed to explain the formation of the cell structure. Wilson suggested that this morphological instability is formed by the intersection of the etched surface with point defect clusters that are created in the dense collision cascade, and the combined effects of sputter etching, ion reflection and redeposition.¹ Others proposed a model of plastic or viscous flow.⁴ However, the un-

derstanding of neither mechanism is clear. With the capability of AFM to measure surface roughness, complementary information on the formation mechanisms of cell structures can be given as below due to the representation of roughness as the dynamic response of the surface to ion bombardment. (For discussion of future scanning tunneling microscopy (STM) study of low dose self implanted Ge samples, see Ref. 16.)

Even though AFM images show distortions from true sample morphology as discussed above, contact mode AFM is a good and convenient tool for the measurement of roughness because it can quantitatively measure the height profile of a surface with very high resolution. Root-mean-square roughness (R_q) is often used because of all the roughness

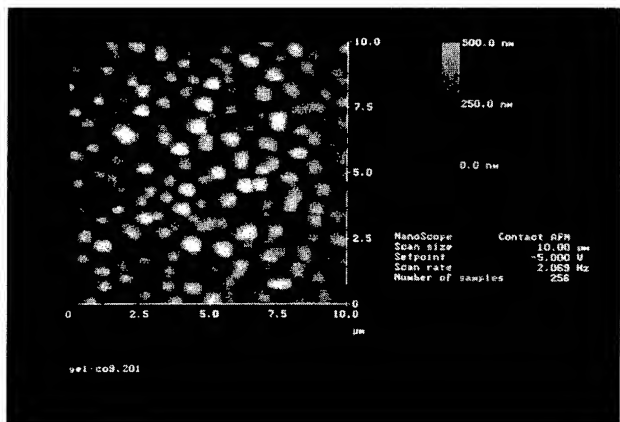


FIG. 6. A contact mode AFM image of polycrystalline granular structures for the highest beam current density and highest dose implantation (dose: 5×10^{17} ions/cm²; accelerating voltage: 70 kV; beam current density: 210 μ A/cm²).

parameters it is the least affected by tip induced distortion.¹⁰ Figure 5(a) shows the evolution of the roughness (R_q over an area of $1 \mu\text{m} \times 1 \mu\text{m}$) with ion dose at two accelerating voltages during Co implantation with a low beam current density of $15 \mu\text{A}/\text{cm}^2$ and an implant temperature (due to beam current self-heating) estimated to be below 200°C . R_q increased exponentially with ion dose from 10^{16} to 10^{17} cm⁻², similar to the relationship of mean cell diameter versus ion dose for 50 keV self-implanted Ge samples.¹ It is also seen that at a lower accelerating voltage and, therefore, lower ion energy the surfaces are rougher. This is shown more clearly in Fig. 5(b). There was a reduction in R_q with increasing accelerating voltage at three doses. This is reasonable because the lower the ion energy, the less the projected range and therefore the more the surface topography is affected. This also fits well with the expected decrease in sputtering with increasing ion energy. It has been observed that the implanted structure is temperature dependent. Figure 5(c) shows the plot of R_q versus beam current density at an accelerating voltage of 70 kV and a dose of 5×10^{16} cm⁻². R_q increases quickly at beam current densities from 15 to 30 $\mu\text{A}/\text{cm}^2$ and then saturates from 30 to 150 $\mu\text{A}/\text{cm}^2$. The temperature corresponding to a beam current density of 30 – 150 $\mu\text{A}/\text{cm}^2$ is $350 - 600^\circ\text{C}$. This dependence is believed to be due to competition between amorphization and radiation-enhanced annealing. At lower temperature, below about 400°C , the implanted surface layer remains amorphous. Between 250 and 350°C , surface diffusion and viscous flow play a more and more important role in surface morphology until an equilibrium state is reached at the amorphous/crystal transition at 400°C . With the highest beam current density ($210 \mu\text{A}/\text{cm}^2$) used in this experiment, the temperature (estimated to be $800 \pm 50^\circ\text{C}$) is high enough to self-anneal the implanted layer into a polycrystalline granular structure as

shown in Fig. 6 for a specimen implanted to a dose of 5×10^{17} cm⁻². We surmise that a stable compound, Co_xGe_y , was formed.

IV. CONCLUSIONS

In summary, cellular structures of Co implanted Ge(111) surfaces were observed by contact mode and tapping mode atomic force microscopy, and are in agreement with earlier SEM/TEM results. The image distortions observed are attributed to a combination of the effects of tip shape and the softness and stickiness of the implanted surface layers. The variation of roughness (R_q) with ion dose, ion energy (accelerating voltage) and temperature (beam current density) was also determined.

ACKNOWLEDGMENTS

This work was supported in part by the Research Grants Committee of the Hong Kong University Grant Council. The authors would like to acknowledge Q. C. Peng for several useful discussions. One of the authors (W.Y.C.) was supported by the Postdoctoral Fellow Scheme of The Chinese University of Hong Kong.

- ¹I. H. Wilson, J. Appl. Phys. **53**, 1698 (1982).
- ²B. R. Appleton, O. W. Holland, J. Narayan, O. E. Schow III, J. S. Williams, and E. Lawson, Appl. Phys. Lett. **41**, 711 (1982).
- ³O. W. Holland, B. R. Appleton, and J. Narayan, J. Appl. Phys. **54**, 2295 (1983).
- ⁴L. M. Wang and R. C. Birtcher, Appl. Phys. Lett. **55**, 2494 (1989); Philos. Mag. A **64**, 1209 (1991), and references therein.
- ⁵R. C. Birtcher, Mater. Res. Soc. Symp. Proc. **279**, 129 (1993).
- ⁶I. H. Wilson, Y. J. Chen, J. B. Xu, R. A. B. Devine, and C. Jeynes, Surf. Interface Anal. **24**, 881 (1996).
- ⁷Q. Peng, S. P. Wong, J. B. Xu, and I. H. Wilson, Mater. Res. Soc. Symp. Proc. **396**, 763 (1996); another example is our recent AFM observation of iron (Fe) implanted Si(100) surfaces.
- ⁸J. E. Griffith, D. A. Grigg, M. J. Vasile, P. E. Russell, and E. A. Fitzgerald, J. Vac. Sci. Technol. B **9**, 3596 (1991), and references therein.
- ⁹P. Grutter, W. Zimmermann-Edling, and D. Brodbeck, Appl. Phys. Lett. **60**, 2741 (1992).
- ¹⁰K. L. Westra and D. J. Thomson, J. Vac. Sci. Technol. B **12**, 3176 (1994); *ibid.*, **13**, 344 (1995); Thin Solid Films **257**, 15 (1995).
- ¹¹G. Kaup, J. Schmeyer, U. Pogodda, M. Haak, T. Marquardt, and M. Plagmann, Thin Solid Films **264**, 205 (1995).
- ¹²L. Hellemans, K. Waeyaert, F. Hennau, L. Stockman, II. Heyvaet, and C. Van Haesendonck, J. Vac. Sci. Technol. B **9**, 1309 (1991).
- ¹³F. Atamny and A. Baiker, Surf. Sci. **323**, L314 (1995).
- ¹⁴R. C. Birtcher, M. H. Grimsditch, and L. E. McNeil, Phys. Rev. B **50**, 8990 (1994).
- ¹⁵Y. J. Chen, W. Y. Cheung, I. H. Wilson, J. B. Xu, and S. P. Wong (unpublished).
- ¹⁶For high dose implantation, we used an air scanning tunneling microscope (STM) to observe the implanted samples and found that there was an increase of electrical conductivity and resistance to oxidation compared with unimplanted surfaces on which the tunneling current was unstable when the tip was stationary and a very noisy trace was obtained upon scanning. This is similar to cobalt and iron implanted silicon samples; see, for example, A. A. Bukharaev, F. F. Gubydullin, A. V. Nazarov, and N. V. Berdunov, Phys. Status Solid A **131**, 79 (1992); for low dose implantation, ultrahigh vacuum (UHV) STM is required to avoid the effects of oxidation.

Nanometer table-top proximity x-ray lithography with liquid-target laser-plasma source

L. Malmqvist,^{a)} A. L. Bogdanov,^{b)} L. Montelius,^{b)} and H. M. Hertz^{a),c)}

Department of Physics, Lund Institute of Technology, P.O. Box 118, S-221 00 Lund, Sweden

(Received 25 October 1996; accepted 4 April 1997)

A compact laser-plasma proximity x-ray lithography system suitable for laboratory-scale low-volume nanometer patterning is presented. The laser-plasma source, which is based on a fluorocarbon liquid-jet target, generates high-brightness $\lambda = 1.2\text{--}1.7$ nm x-ray emission with only negligible debris production. The Au/SiN_x x-ray mask is fabricated by employing ion milling and a high-contrast e-beam resist. With SAL-601 chemically enhanced resist we demonstrate fabrication of high-aspect-ratio, sub-100 nm structures. The exposure time is currently 20 min using a compact 10 Hz, $\lambda = 532$ nm, 70 mJ/pulse mode-locked Nd:YAG laser. However, the regenerative liquid-jet target is designed for operation with future, e.g., 1000 Hz, lasers resulting in projected exposure times of ~ 10 s. © 1997 American Vacuum Society. [S0734-211X(97)01004-4]

I. INTRODUCTION

Proximity x-ray lithography shows promise of becoming a suitable technique for the fabrication of nanometer structures.^{1–5} Pattern transfer with feature sizes of a few tens of nanometers is feasible.² Most current systems are, unfortunately, based either on expensive synchrotron sources^{1,2} having limited accessibility, or conventional low-intensity electron impact sources^{3,4} that result in long exposure times. This has limited the spread of the technology into the nanofabrication community. The laser plasma is an alternative high-brightness x-ray source.⁵ However, previous laser-plasma-based lithography has utilized larger-scale lasers and conventional target systems, such as bulk metal^{6–8} or thin metal films.⁹ Such target systems have a limited operating time between the necessary changes of the target. They do not allow a high laser repetition rate, which results in a limited average x-ray flux. They also suffer from severe debris emission making mask damage a problem. By using microscopic liquid fluorocarbon droplets as the target, the debris problem is reduced by several orders of magnitude.¹⁰ Furthermore, the droplet x-ray source¹¹ is compact, may be operated at high repetition rate, provides nearly 4π steradian access and is practically unlimited in operating time. In the present article we combine a modified version of the droplet x-ray source with a simple method for x-ray mask fabrication, demonstrating laboratory-scale lithography of high-aspect ratio sub-100 nm structures. This compact system has potential for short-exposure times and is particularly appropriate for low-volume laboratory-scale applications of nanolithography.

II. EXPERIMENTAL ARRANGEMENT

The operation of the original table-top fluorocarbon-droplet source¹⁰ was modified by employing a microscopic liquid jet for enhanced long-term stability and simpler tem-

poral synchronization.¹² The experimental arrangement of the lithography system is shown in Fig. 1. A fluorocarbon liquid is forced through an ~ 10 μm capillary glass nozzle by a pressure of ~ 40 bar into a $\sim 10^{-4}$ mbar vacuum chamber. This produces a microscopic fluorocarbon liquid jet that is a few mm long. The diameter is ~ 15 μm and the velocity is ~ 50 m/s. The jet is irradiated from the side by the focused beam from a 10 Hz mode-locked Nd:YAG laser (Continuum PY61C-10). The laser generates 120 ps, $\lambda = 1064$ nm, 140 mJ pulses. By the frequency conversion and mixing in three nonlinear crystals this is converted into a 3 mJ, $\lambda = 355$ nm prepulse which hits the liquid-jet target 7 ns before the 70 mJ, $\lambda = 532$ nm main pulse.¹³ The laser pulses are converted into x rays in the $\lambda \approx 1.2\text{--}1.7$ nm spectral range with a conversion efficiency of approximately 4%, resulting in $\sim 1.8 \times 10^{12}$ ph/sr/pulse. The conversion efficiency is slightly smaller than that in Ref. 10, which may be attributed to the longer laser pulses used in the present experiment. The $\lambda < 1.7$ nm x rays are mainly due to F IX and F X emission. The full width at half-maximum (FWHM) diameter of the x-ray emitting plasma was measured to be 30 μm with a pinhole camera. The small size of the plasma as well as negligible debris make it possible to use a small distance (40 mm) between the plasma and the mask/substrate unit, while still maintaining an acceptable penumbral blur. This increases the x-ray intensity at the substrate. The target liquid not used for x-ray generation is collected in a liquid nitrogen trap.

The lithography exposure chamber is equipped with a loadlock system for rapid exchange of mask/substrate units while the liquid-jet x-ray source is operating in vacuum. Three spring-loaded screws on a stainless steel chuck press the mask against the resist-coated Si substrate via a 13 μm polyimide spacer. The mask/substrate unit is then inserted into the exposure chamber so that the substrate surface is positioned ~ 40 mm from the plasma.

The x-ray mask was manufactured on a 250 nm thick SiN_x membrane¹⁴ covered by an absorbing layer of 185 nm Au and 3 nm NiCr. Patterning of the mask's absorbing metal

^{a)}Division of Atomic Physics.

^{b)}Division of Solid State Physics.

^{c)}Electronic mail: hans.hertz@fysik.lth.se

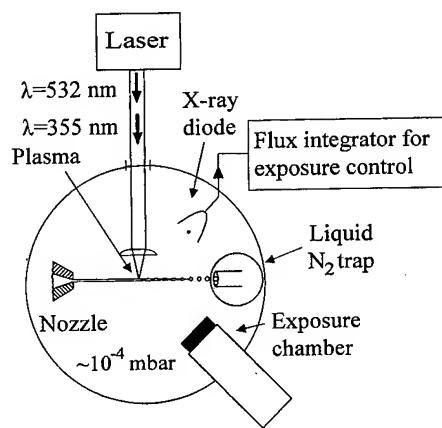


FIG. 1. Experimental arrangement for table-top nanometer x-ray lithography using liquid-jet laser-plasma source.

layers was performed by Ar^+ ion milling. The resist mask for the ion milling was prepared by e-beam exposure of the high-contrast positive resist Zeonrex ZEP-520.¹⁵ To minimize proximity effects, 50 keV electrons were used for the e-beam writing (JEOL JBX-5DII). Electrons with this energy were found to pass through the Au/NiCr/SiN_x substrate without any significant backward scattering. This resulted in a considerable reduction of the long-range proximity effect, whereas the short-range effect was noticeable at distances shorter than the thickness of the Au film. The resist development conditions were chosen to obtain the highest possible contrast ($\gamma=8$ for a 0.25 μm thick film), which made it possible to obtain resist profiles with better than 2:1 height-to-width ratio. After the e-beam lithography, the gold was patterned by Ar^+ ion etch through the resist mask. During the etching the mask was rotated and the axis of rotation was tilted approximately 5° to the direction of the ion beam. This method allowed us to control the height of the edge fences resulting from redeposition of etched material onto the resist walls. The smallest resolved Au features of the mask were below 100 nm in size. Structures larger than 200 nm exhibit nearly uniform 185 nm thickness and very sharp edges. Figure 2 shows a part of the mask. The pattern on the mask was chosen to allow critical evaluation of the x-ray lithography method. It included similar features to those based on 100, 250, and 500 nm design rules. Due to the proximity effect

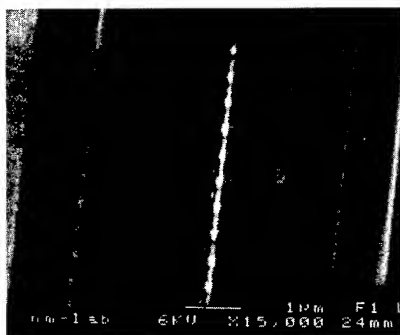


FIG. 2. Scanning electron micrograph of a section of the Au/SiN_x x-ray mask.

mentioned above, the features with sizes below 200 nm were difficult to reproduce in the resist without active proximity correction.

The 185 nm Au/250 nm SiN_x mask provides an x-ray transmission contrast of about 11:1 for the $\lambda \approx 1.5$ nm laser-plasma emission. To suppress the longer-wavelength emission in the spectrum, which is mainly due to C V and C VI ions, a free-standing metal filter consisting of 130 nm Cu and 1 μm Al is inserted between the mask and the plasma. The total transmission of the $\lambda < 1.7$ nm radiation through the SiN_x mask and filter combination is approximately 40%. The SiN_x membranes employed for the mask were 0.5 mm \times 0.5 mm on a Si substrate.¹⁴ Much larger-area membranes (1–2 μm thick) with extreme strength are routinely manufactured.³ Naturally such membranes can be combined with the system described in this article. If such a membrane would be used for the mask instead of the ones we had available, the metal filters could be omitted. Compared to conventional x-ray lithography around $\lambda = 1\text{--}1.2$ nm, the slightly longer wavelength in our experiments makes mask fabrication simpler due to increased gold absorption, thereby avoiding the need for fabrication of very high aspect ratios in the absorber material.

An x-ray diode covered by 2 μm Al and 260 nm Cu free-standing metal filters is used for on-line determination of the x-ray flux in the $\lambda < 1.7$ nm region. The higher filter thickness was chosen to avoid saturation of the diode. From diode sensitivity data and filter transmission calculations, the x-ray intensity was calculated to be 4 $\mu\text{J}/\text{cm}^2/\text{pulse}$ at the substrate surface. The absolute accuracy of this data is 50%; the uncertainty is due to the filter thickness, the x-ray diode sensitivity data and the relative strength of the spectral lines.¹¹ In order to obtain a controlled x-ray dose in each exposure, the diode signal, which corresponds to the emitted x-ray flux in each shot, was electronically integrated over time. Since the x-ray flux was determined to be approximately uniform in all directions, the total deposited energy in the resist could be accurately estimated.

III. RESULTS AND DISCUSSION

X-ray lithography was performed with the acid-catalyzed, novolak-based negative resist Microposit SAL-601 ER7,¹⁶ which is approximately 20 times more sensitive than PMMA. An approximately 0.6 μm thick film was spin coated onto Si substrates. The resist processing included a 110 $^\circ\text{C}$ postexposure baking step in a convection oven followed by 2 min of development in Microposit MF-322.¹⁶ The sensitivity curve obtained with these exposure/development parameters and with a 22 min postexposure baking time is shown in Fig. 3. After optimization of the process, we used an exposure dose of 45 mJ/cm^2 and a postexposure baking time of 18 min for the best lithography results.

Figure 4(a) shows a scanning electron microscope (SEM) micrograph of a typical lithographic pattern. The exposure time was 20 min. In Fig. 4(b) the pillars in the frame of Fig. 4(a) are displayed at a higher magnification. These high-

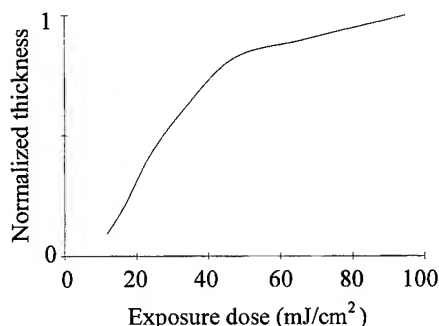


FIG. 3. Sensitivity curve of SAL-601 negative resist with 22 min postexposure baking time.

aspect-ratio pillars have a typical diameter of 70 nm, which is determined mainly by the mask. The wall profiles are vertical or even leaning slightly backward, indicating that this lithography procedure may be suitable for lift-off processing of metal patterns on the nanometer scale. The slight undercut may result from the comparatively high absorption ($\sim 25\%$) of the incident $\lambda \approx 1.2\text{--}1.7$ nm x rays in the resist layer. Thus, a slightly larger fraction of the absorbed energy is deposited in the top layers of the negative resist resulting in a higher density insoluble resist there, while the less exposed lower layers are less crosslinked.

Resolution in proximity x-ray lithography is primarily limited by diffraction and by the photoelectron range.^{2,3} The photoelectron range sets the ultimate limit since diffraction, in principle, can be reduced by reducing the mask-substrate

gap. For wavelengths of the present liquid-jet x-ray source, i.e., $\lambda \approx 1.5$ nm, the maximum range of the photoelectrons is 20–30 nm while most of the energy deposition is due to the short-range (~ 5 nm) Auger electrons.³ Thus, with $\lambda \approx 1.5$ nm, < 20 nm features would be feasible when considering only the photoelectrons.² Shorter wavelength lithography systems will increase the photoelectron range.

The minimum linewidth w due to diffraction may be estimated from

$$w = k_1 \sqrt{\lambda \cdot g}, \quad (1)$$

where g is the gap distance between the mask and the substrate, λ is the wavelength, and k_1 is a factor that depends on the mask and the illumination system. Using $k_1 \approx 0.6$ (cf. Ref. 2) and the parameters used in the present article ($\lambda \approx 1.5$ nm, $g \approx 13$ μm) results in a minimum achievable linewidth of ~ 90 nm. With a 2 μm gap, routinely used in Ref. 3 even for large-diameter masks, this reduces to ~ 35 nm. Finally, it may be noted that previous investigations show that some spatial incoherence in the source is favorable to avoid edge ringing.³ In our experiments we had a penumbral blur δ of ~ 10 nm, resulting in the blur-to-width ratio $\beta = \delta/w$ being approximately 0.1. Studies show that a somewhat larger β (approximately 0.3) provides better contrast.¹⁷ Thus, a larger plasma diameter or smaller plasma-mask distance may improve the current arrangement.

The exposure time is approximately 20 min in the current arrangement, which is based on the 10 Hz laser-plasma source. Although this is reasonable for many low-volume laboratory-scale projects, several applications require shorter exposures times. This is especially true for larger-volume fabrication. Fortunately, the average x-ray flux from our negligible debris laser-plasma source may be increased significantly since the liquid-jet target allows much higher repetition-rate lasers to be used. This is in contrast to many other laser-plasma sources and is due to the fact that fresh target material is supplied in a continuous flow at a rapid rate (jet velocity ~ 50 m/s). Lasers aiming at a 1000 Hz repetition rate, and with suitable pulse energy and pulse width, are currently being developed.¹⁸ With a 1000 Hz laser the exposure time would be on the order of 10 s. Furthermore, the effective throughput may be increased since the nearly 4π steradian geometric access allows several exposure chambers at the source to be operated simultaneously. Finally, it might be of importance that the cost of the target is very low (approximately $\$10^{-7}$ /pulse for a 1000 Hz system¹⁰) due to the small liquid flow and the intrinsic accurate forming and positioning of the target material.

IV. SUMMARY

We have demonstrated a compact laser-plasma proximity x-ray lithography system appropriate for low-volume nanometer patterning. It combines a Au/SiN_x mask and a sensitive negative x-ray resist with a fluorocarbon liquid-jet target laser-plasma x-ray source with negligible debris emission. The lithography system with its 10 Hz current source is already suitable for low-volume small-scale appli-

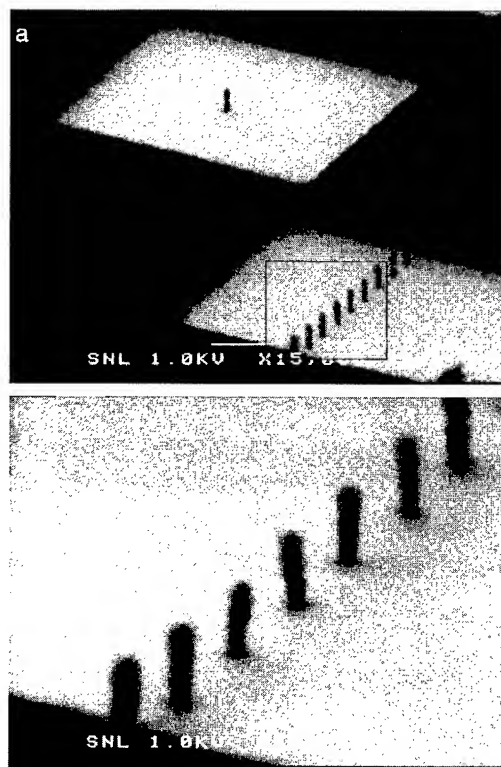


FIG. 4. Electron micrographs at low (a) and high (b) magnification of high-aspect-ratio 70 nm diameter pillars fabricated in SAL-601 negative resist.

cations in, for example, research environments, and has the potential for larger-scale applications using higher-repetition-rate lasers.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the valuable discussions with and assistance from Lars Rymell, Magnus Berglund, Derrick Mancini, Terje Rye (Siemens-Elema, Stockholm), and Sune Svanberg. This work was financed by the Swedish Board for Industrial and Technical Development, the Swedish Natural Science Research Council, the Wallenberg Foundation and the Nanometer Structure Consortium.

¹A. Heuberger, *J. Vac. Sci. Technol. B* **6**, 107 (1988).

²M. Kahn, L. Mohammed, J. Xiao, L. Ocola, and F. Cerrina, *J. Vac. Sci. Technol. B* **12**, 3930 (1994).

³H. I. Smith and M. L. Schattenburg, *IBM J. Res. Dev.* **37**, 319 (1993).

⁴W. Chu, H. I. Smith, and M. L. Schattenburg, *Appl. Phys. Lett.* **59**, 1641 (1991).

⁵J. R. Maldonado, *Proc. SPIE* **2523**, 2 (1995).

⁶M. Kuehne and H.-C. Petzold, *Appl. Opt.* **27**, 3926 (1988).

⁷F. Bijkerk, E. Louis, M. J. van der Wiel, E. C. I. Turcu, G. J. Tallents, and D. Batani, *J. X-Ray Sci. Technol.* **3**, 133 (1992).

⁸M. L. Schattenburg, K. Early, Y.-C. Ku, W. Chu, M. I. Shepard, S.-C. The, H. I. Smith, D. W. Peters, R. D. Frankel, D. R. Kelly, and J. P. Drumheller, *J. Vac. Sci. Technol. B* **8**, 1604 (1990).

⁹C. M. Reeves, I. C. E. Turcu, P. D. Prewett, A. M. Gundlach, J. T. Stevenson, A. J. Walton, A. W. S. Ross, R. A. Lawes, P. Anastasi, R. Burge, and P. Mitchell, *Electron. Lett.* **31**, 2218 (1995).

¹⁰L. Malmqvist, L. Rymell, and H. M. Hertz, *Appl. Phys. Lett.* **68**, 2627 (1996).

¹¹L. Rymell and H. M. Hertz, *Opt. Commun.* **103**, 105 (1993).

¹²L. Malmqvist, L. Rymell, M. Berglund, and H. M. Hertz, *Rev. Sci. Instrum.* **67**, 4150 (1996).

¹³M. Berglund, L. Rymell, and H. M. Hertz, *Appl. Phys. Lett.* **69**, 1683 (1996).

¹⁴FaSTec Ltd., Great Britain.

¹⁵Nippon Zeon Co. Ltd., Japan.

¹⁶Shipley Inc., Norrköping, Sweden.

¹⁷S. D. Hector, M. L. Schattenburg, E. H. Anderson, W. Chu, V. V. Wong, and H. I. Smith, *J. Vac. Sci. Technol. B* **10**, 3164 (1992).

¹⁸JAMAR Inc., San Diego, CA.

Low-energy focused-ion-beam exposure characteristics of an amorphous $\text{Se}_{75}\text{Ge}_{25}$ resist

Hyun-Yong Lee and Hong-Bay Chung^{a)}

Department of Electronic Materials Engineering and Institute of New Technology, Kwangwoon University, 447-1, Wolgyedong, Nowonku, Seoul 139-701, Korea

(Received 6 March 1996; accepted 4 April 1997)

Ion-beam exposure and patterning characteristics of an amorphous $\text{Se}_{75}\text{Ge}_{25}$ thin film which acts as a positive resist in focused-ion-beam (FIB) lithography have been studied using low-energy Ga^+ ion sources below 30 keV. The $a\text{-Se}_{75}\text{Ge}_{25}$ resist exposed to 10 keV defocused- Ga^+ ion beam with the dose of 5.0×10^{14} to 9.3×10^{15} ions/cm² results in increasing the optical absorption, which was also observed in the film exposed to an optical dose of 4.5×10^{20} photons/cm². The absorption edge shift of 0.3 eV for the resist exposed to a dose of 9.3×10^{15} ions/cm² at 10 keV is about twice that of photo exposure. These large shifts could be estimated as due to an increase in disorder, considering a decrease in the slope of the Urbach tail and a broad pattern of x-ray diffraction. For the exposure of a 30 keV Ga^+ ion beam and above a dose of 1.4×10^{15} ions/cm², a 590-Å-thick resist film is completely etched by dipping for 10 s in 1:1:3 $\text{HNO}_3\text{:HCl:H}_2\text{O}$ solution (25 °C), and then the etching rate is about 60 Å/s. As the incident energy increases from 10 to 30 keV, the threshold dose decreases from 4.0×10^{15} to 1.4×10^{15} ions/cm² and then the imaging contrasts appear to be about 0.5 and 2.5, respectively. The decrease of the threshold dose with increasing the exposure energy is evidence that a predominant factor in the FIB exposure characteristics is the energy transfer rather than the implanted ions themselves. When 30 keV Ga^+ FIB exposure with a multiscan diameter of 0.2 μm and the above-mentioned development are employed, a resist pattern with a linewidth of about 0.225 μm is obtained. © 1997 American Vacuum Society. [S0734-211X(97)01104-9]

I. INTRODUCTION

In recent years, focused-ion-beam (FIB) technology has advanced rapidly with the requirement for scaling down electronic devices to the subquarter micrometer region.¹ In particular, focused-ion-beam lithography (FIBL) is evaluated as a novel vacuum technique available to generate original patterns (or masks), with the sub-0.1 μm linewidth, for x-ray lithography (or ion-projection lithography).^{2,3} In FIBL, using conventional polymer-resist materials, one of the main problems is that the ion penetration through half a micron of a polymer, essential for conventional wet-chemical development, requires either lower mass ions or medium mass ions having 50–200 keV of incident energy. High-energy ions penetrate deep into the sample and yield relatively high resolution, sharp resist profiles. However, when high-energy FIB is used for direct exposure of the resist over a substrate without a mask, damages by unwanted implantation into the substrate may be an issue.⁴ The possible ways to avoid this are to select the resist thickness capable of minimizing ion penetration into a substrate and to use an inorganic resist with large ion-stopping power or low-energy FIBL.⁵

Amorphous germanium selenide thin films have been investigated as inorganic resists for microlithography techniques using a variety of exposure sources, such as visible, UV, and deep UV lights, as well as electron beam and ion beam.^{6,7} These materials offer an attractive potential alternative to conventional polymer resists because of a large number of advantages. Especially, these films for FIB exposure

are dense enough to allow fabrication of ultrafine patterns, have greater ion-stopping power to protect a substrate from ion-induced damage than SiO_2 and conventional polymer resists which are basically carbon compounds, and offer surface imaging known to be an effective method of obtaining high imaging contrast.^{8,9} We have recently presented the results of a three-dimensional Monte Carlo (3D MC) simulation for Ga^+ ion penetration in an amorphous germanium selenide thin film and the development profiles of FIBL using the simplified exposure and development models.^{10–12}

In this article, we present ion-induced changes, exposure response characteristics, and development patterns of the $a\text{-Se}_{75}\text{Ge}_{25}$ resist by a low-energy Ga^+ FIB exposure which has been conventionally utilized for mask repairing. Then the ion range parameters and the minimum resist thickness are determined by Lindhard, Scharff, and Schiott (LSS)-based calculation^{13,14} and MC simulation.^{11,12} The optical absorption edges of the chalcogenides are characterized by an absorption constant α that rises exponentially with increasing photon energy up to a value of α of in the range $10^3\text{--}10^4\text{ cm}^{-1}$. Some chalcogenides—especially, Se or S-based—materials exhibit an absorption tail, a so-called Urbach tail, for photon energy ($\hbar\omega$) just below the optical energy gap E_{op} and obey the absorption properties expressed as $\alpha = \alpha_0 \exp[(\hbar\omega - E_{\text{op}})/F]$ and $\alpha\hbar\omega = B(\hbar\omega - E_{\text{op}})^2$ for energy above E_{op} (i.e., the extended energy region) in which α_0 is a material parameter, $1/F$ is exponential slope in extended region, and constant B lies in the range $10^5\text{--}10^6\text{ cm}^{-1}\text{ eV}^{-1}$.^{15,16} These absorption properties around the Urbach tail have been studied in relation to the

^{a)}Electronic mail: hychung@daisy.kwangwoon.ac.kr

TABLE I. The calculated parameters and the minimum resist thickness for the combination of α -Se₇₅Ge₂₅ resist and low-energy Ga⁺ FIB. LSS-based results are compared with MC results, where R , R_p , ΔR_p , and ΔR_\perp are total ion path, projected ion range, projected range straggling, and lateral range straggling, respectively. In the case of MC simulation comprising the backscattering effect of ion projectile, the projected straggling is separated into ΔR_{p1} and ΔR_{p2} due to asymmetrical concentration distribution, where ΔR_{p1} and ΔR_{p2} are for right and left side in the peak concentration $N(R_p)$, respectively, and ΔR_\perp values are in good agreement with those of ΔR_{p2} for all energy ranges. The Z_{\min} of an α -Se₇₅Ge₂₅ resist for 30 keV Ga⁺ FIB lies in the range of 546–590 Å for MC simulation and has 582 Å for LSS-based calculation.

Incident energy								
E	(keV)	R (Å)	R_p (Å)	ΔR_\perp (Å) ^a	ΔR_p (Å)	$\Delta R_\perp / \Delta R_p$	Z_{\min} (Å)	$N(R_p)^b$ (10 ²⁰ ions/cm ³)
10	LSS	137	101	40	43	0.93	273	9.1
	MC	121	88	42	$\Delta R_{p1}:50$ $\Delta R_{p2}:42$	0.84 1.00	288 256	8.7
20	LSS	215	158	65	70	0.93	438	5.6
	MC	192	141	70	$\Delta R_{p1}:81$ $\Delta R_{p2}:69$	0.86 1.01	465 417	5.0
30	LSS	286	210	86	93	0.93	582	4.3
	MC	259	190	91	$\Delta R_{p1}:100$ $\Delta R_{p2}:89$	0.91 1.02	590 546	4.0

^a ΔR_\perp for LSS-based data is obtained by using the Schiott empirical formula (Ref. 22).

^bGa⁺ ion dose is 1×10^{15} ions/cm².

photostructural change or the photodarkening in amorphous-chalcogenide semiconductors.¹⁷ Two exponential slopes, i.e., the slope of Urbach tail U and $1/F$, are obtained on a plot of α vs $\hbar\omega$ and then can be changed according to several treatment conditions. The exponential slope reduction in photo- or ion-exposed films has been reported due to the increase of randomness in atomic configuration.^{18–20} Therefore, the evaluation of absorption changes in amorphous SeGe chalcogenide film by ion exposure could be interpreted in relation to microstructural rearrangements within the short range of the SeGe network, that is the atomic configuration of the system reaches a new metastable state. Due to changes in this short range order, the chemical bonding may be affected, which results in enhanced chemical dissolution in ion-exposed films.⁶

II. CALCULATION AND EXPERIMENT

The use of thinner ion-beam resist leads to the fine pattern definition. For practical use, however, it is necessary to select an appropriate resist thickness to absorb the ion-beam energy sufficiently and to minimize a substrate damage at the same time. We obtained the minimum resist thickness Z_{\min} as well as the ion-range parameters, such as R_p , ΔR_p , ΔR_\perp , calculated using MC simulation^{10–12} and LSS-based theory^{13,14} for FIBL with variable incident energies. It will be adequate for the general case to consider the transmission coefficient of impinging ions T for estimating Z_{\min} . Assuming a Gaussian distribution of implanted ions, Z_{\min} can be expressed as $Z_{\min} \cong R_p + 4\Delta R_p$ for $T = 10^{-4}$.²¹ As an example of this definition, when ion beam impinges on the Z_{\min} -thick resist/Si substrate with a dose of 1×10^{16} ions/cm², the number of ions implanted into the substrate is about 1×10^{12} ions/cm², which corresponds to about 0.1 at. % considering that the atomic density on the (100) surface of Si substrate is 6.7×10^{14} atoms/cm². A den-

sity of α -Se₇₅Ge₂₅ is dependent on the fabrication and annealing processes, and for the most case lies between 4.3 and 4.45 g/cm³ without annealing.²² In our calculation, the densities of α -Se₇₅Ge₂₅, α -Se, and α -Ge are 4.42, 4.28, and 4.85 g/cm³, respectively, which correspond to 3.39, 3.26, and 4.02×10^{22} atoms/cm³ of the atomic density.²³ The calculated parameters for the combination of α -Se₇₅Ge₂₅ resist and low-energy Ga⁺ FIB are summarized in Table I.

An amorphous Se₇₅Ge₂₅ thin film was prepared by thermal evaporation of the bulk in vacuum. A Se₇₅Ge₂₅ bulk glass was prepared by a conventional melt quenching technique. The constituent elements Se and Ge weighed in the appropriate ratio were sealed in an evacuated quartz ampoule, which was then placed in a furnace and heated at 220, 650, and 1000 °C for 2, 2, and 24 h, respectively.

The ampoule was constantly stirred during these periods to achieve complete homogenization of the constituents in the melt and quenched in water. Films were prepared by thermal evaporation at a deposition rate of about 5 Å/s on Corning glass-7059 (density: 2.76 g/cm³) and p -type Si (100) kept at room temperature in vacuum $\sim 1 \times 10^{-5}$ Torr. Films were deposited in three thickness types: about 2000 Å (the conventional photoresist thickness) on glass to observe ion-induced properties, about 800 Å on Si wafer to evaluate the etching rate, Z_{\min} (Table I), on Si wafer to obtain the exposure response and pattern characteristics. The composition of the films was checked by Auger electron spectroscopy. From the results, a resist composition evaluated to be consistent with that of bulk in the range of accuracy of 95%–99% (i.e., Se:Ge=76:24 at. %).

Low-energy ion-induced properties were observed while exposing 10 keV defocused ion beam with the dose of 5×10^{13} – 3×10^{16} ions/cm². The gap spacing between a Ga⁺ LMIS source and an extractor was 0.2 cm and a target lay below 2 cm from an extractor. A UV-visible spectrophotom-

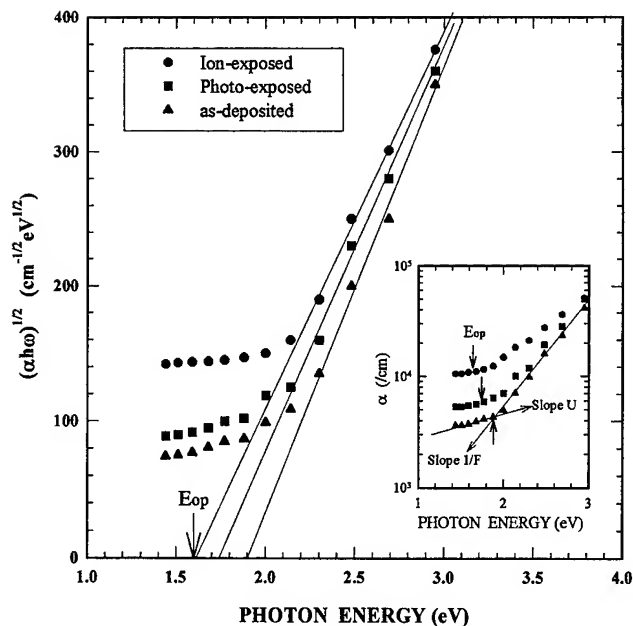


FIG. 1. The optical energy gap E_{op} and the optical absorption constant α of $a\text{-Se}_{75}\text{Ge}_{25}$ resist with the exposure of defocused Ga^+ ion beam (ion dose of 9.3×10^{15} ions/cm² with 10 keV) and optical radiation (optical dose of 4.5×10^{20} photons/cm² with an energy above E_{op}).

eter (Shimadzu, UV-240) and a N&K analyzer (NKT 1200) were used to record the ion-beam-induced absorption edge shifts in exposed films. The phase nature of exposed and unexposed films was evaluated by x-ray diffraction (XRD) measurement. A low-energy FIB machine (SMI-8100, Seiko Instrum., Inc.) which is generally utilized for mask repairing was the exposure tool for obtaining exposure and development characteristics. The incident energies were 10, 20, and 30 keV and then the exposure dose was controlled by changing the exposure current and time. Development after the exposure was performed by chemical etching in a $\text{HNO}_3 + \text{HCl} + \text{H}_2\text{O}$ (1:1:3) solution at 25 °C. The etching method employed was dipping for a few seconds. The lithographic patterns were obtained by the above-mentioned development for 10 s after 30 keV FIB multiscanning with squares of dimension $50 \times 0.2 \mu\text{m}$ —the dose of 1.4 and 6.5×10^{15} ions/cm². The step height between exposed and unexposed resist were measured using Nanoscope (Anelva) and profilometer, and the lithographic pattern of the developed $a\text{-Se}_{75}\text{Ge}_{25}$ resist was observed with a scanning electron microscope (SEM).

III. RESULTS AND DISCUSSION

The optical absorption constant as a function of wavelength, $\alpha(\lambda)$, can be obtained using the relation $\alpha(\lambda) = \ln[T_{op}(\lambda)/z]$ from the optical transmittance T_{op} or $\alpha(\lambda) = 4\pi k/\lambda$ from the extinction coefficient k . When a 10 keV, defocused Ga^+ ion beam is exposed, the variation of the optical energy gap E_{op} and the absorption constant of an $a\text{-Se}_{75}\text{Ge}_{25}$ resist is shown in Fig. 1, which contains the result induced by optical radiation.

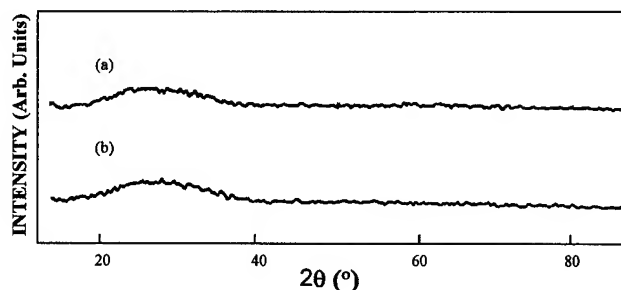


FIG. 2. XRD patterns of exposed and unexposed $a\text{-Se}_{75}\text{Ge}_{25}$ resists. (a) As-deposited; (b) 10 keV Ga^+ ion exposed (8×10^{15} ions/cm²).

The $a\text{-Se}_{75}\text{Ge}_{25}$ resist exposed to a Ga^+ ion dose of 9.3×10^{15} ions/cm², which corresponds to 13.4 J/cm^2 , shows an increase in optical absorption. This shift in the absorption edge toward longer wavelength, called the darkening effect,¹⁷ is also observed in the resist exposed to an optical dose of 4.5×10^{20} photons/cm², which corresponds to 144 J/cm^2 as the result of illuminating for 150 min using the blue-pass-filtered 400 W mercury lamp of a surface intensity 16 mW/cm^2 . Also, the E_{op} obtained by extrapolating $(\alpha\hbar\omega)^{1/2}$ for photon energy ($\hbar\omega$), as shown in Fig. 1, is about 1.9 eV for the as-deposited $a\text{-Se}_{75}\text{Ge}_{25}$ resist and about 1.6 eV for the ion-induced resist. The value of ΔE_{op} ($\approx 0.3 \text{ eV}$) for the ion-induced resist is about twice that obtained for the photoinduced resist and larger than that ($\approx 0.2 \text{ eV}$) reported by Chopra *et al.*,²⁴ which have utilized 50 keV He^+ ion beam with a dose of $\sim 10^{16}$ ions/cm². Considering that the decrease of exponential slopes U and $1/F$ shown in the inner part of Fig. 1 is due to the energy (photon, electron, or ion)-enhanced randomness,^{6,13,14,18-20} the amount of the reduction is thought to contribute to the enhancement of the etch selectivity between exposed and unexposed film. Figure 2 shows XRD patterns for both as-deposited film and 10 keV Ga^+ ion exposed film (8.0×10^{15} ions/cm²), which take the broad shape of amorphous phase for the whole ranges.

A hill in the range $2\theta = 20^\circ - 35^\circ$ relates to Se-Ge bonding and is distributed more extensively in the case of ion-exposed film. The variation of ion-induced properties as a function of the exposed dose of $5.0 \times 10^{14} - 3.0 \times 10^{16}$ ions/cm² is summarized in Table II, where a subscript ° means as-deposited resist.

The values, such as E_{op} , B , $1/F$, and U decrease and the α increases with increasing exposed dose. The primary factor on these variations could be estimated as due to the increase of disorder by the collisions (electronic and nuclear stopping) of energetic ions in $a\text{-Se}_{75}\text{Ge}_{25}$ resist, and the effect of the implanted ion concentration should be also considered. The value B corresponding to the square of the slope in the outer graphs of Fig. 1 is decreased slowly from 1.1×10^5 to $9.0 \times 10^4 \text{ cm}^{-1} \text{ eV}^{-1}$, particularly for the increase of exposed dose from 5.0×10^{14} to 9.3×10^{15} ions/cm², while the decrease of approximately two orders occurs for increasing a dose to 3.0×10^{16} ions/cm². The critical dose for an incident Ga^+ ion energy of 10 keV is on the order of 10^{16} ions/cm².

TABLE II. The variation of ion-induced properties as a function of the exposed dose at 10 keV.

Dose at 10 keV (ions/cm ²)	E_{op} (eV) $E_0=1.90$	B (cm ⁻¹ eV ⁻¹) $B_0=1.15 \times 10^5$	α (cm ⁻¹) at E_{op} $\alpha_0=4.1 \times 10^3$	Extended region ($>E_{op}$)		Edge region ($<E_{op}$)	
				$1/F$ (eV ⁻¹)	$\Delta(1/F)/1/F_0$ 100(%) $1/F_0=2.385$	U (eV ⁻¹)	$\Delta U/U_0$ 100(%) $U_0=0.461$
5.0×10^{14}	1.86	1.1×10^5	4.5×10^3	2.150	9.90	0.451	2.2
1.4×10^{15}	1.80	9.8×10^4	6.0×10^3	2.030	14.92	0.260	43.6
4.0×10^{15}	1.62	9.1×10^4	1.1×10^4	1.287	46.06	0.100	78.3
9.3×10^{15}	1.60	9.0×10^4	1.1×10^4	1.280	46.40	0.086	81.3
3.0×10^{16}	0.80	8.0×10^2	2.6×10^5

and the variation of optical properties by FIB above the critical dose can be thought of as due to the implanted ions themselves rather than the deposited energy.

The selective etching characteristics of 800-Å-thick α -Se₇₅Ge₂₅ resist exposed to a Ga⁺ dose of 1.4×10^{15} ions/cm² at 30 keV, after development in an acid solution of 1:1:3 HNO₃:HCl:H₂O is shown in Fig. 3.

The etching rate of unexposed area R_0 is approximately 2.0 Å/s and that of exposed area R is separated into two cases; $R \sim 60$ Å/s for 0–10 s and $R \sim 9$ Å/s up to 20 s. The depth etched for 10 s is about 600 Å and this is almost consistent with the minimum thickness Z_{min} , shown in Table I. Figure 4 shows typical exposure response curves of Z_{min} -thick α -Se₇₅Ge₂₅ resists exposed to 10, 30, and 80 keV Ga⁺ FIB.

As seen from these curves, sensitivity S , i.e., the threshold ion dose necessary to completely etch out resist films for a given development condition, is about 4.0×10^{15} and 1.4×10^{15} ions/cm² for the incident energies of 10 and 30 keV, respectively. The S of α -Se₇₅Ge₂₅ as a positive ion resist (~ 7 J/cm² at 30 keV) is better than its S as a photo ($\sim 1.6 \times 10^{19}$ photons/cm² ~ 10 J/cm²) or electron beam ($\sim 6 \times 10^{15}$ electrons/cm² ~ 20 J/cm² at 20 keV) positive resist.²⁵ The expose response property for a relatively high energy 80 keV is shown by the filled triangle

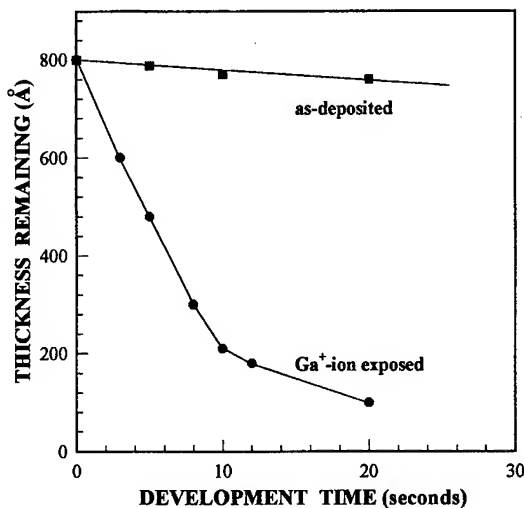


FIG. 3. The etching characteristic of α -Se₇₅Ge₂₅ thin film exposed to a dose of 1.4×10^{15} ions/cm² at 30 keV and developed in 1:1:3 HNO₃:HCl:H₂O solution at room temperature.

($Z_{min}=1180$ Å)¹² in Fig. 4 and the sensitivity is expected to be below 1.0×10^{15} ions/cm². As the incident energy increases from 10 to 30 keV, the threshold dose reduces about three times and the slope around the threshold dose, i.e., the imaging contrast γ , becomes more steep— γ increases from 0.5 to 2.5. The mechanism of a low-energy Ga⁺ FIB lithography using α -Se₇₅Ge₂₅ can be related to the Ga ion distribution in the resist.¹⁴ As shown in Table I, both R_p and ΔR_p approximately double and $N(R_p)$ decreases to about half when the incident ion energy is increased from 10 to 30 keV. If the reaction of implanted ions themselves is a predominant factor in FIBL, the threshold dose should be increased and the contrast should also become less steep with increasing incident energy.²⁶ As the above-mentioned discussions about Fig. 1, Table II, and Fig. 4, FIB lithography in our present work, depends primarily on the energy density deposited by the collisions of energetic ions. Moreover, by taking into account that the ion stopping power for the combination of the α -Se₇₅Ge₂₅ resist and Ga⁺ ions exhibits a maximum at about 70 keV, we can postulate that S does not

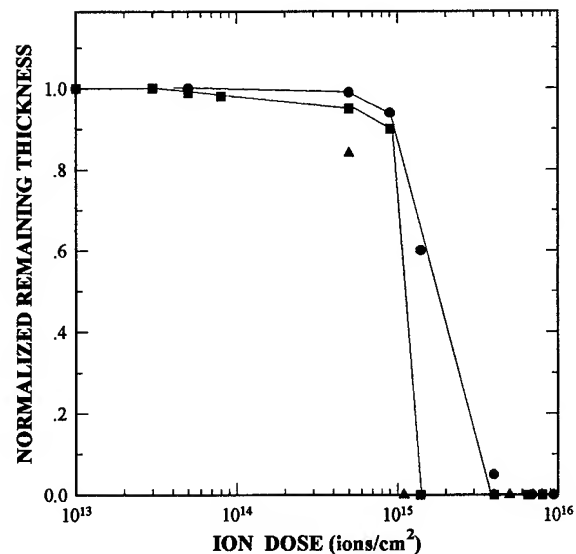


FIG. 4. Exposure response characteristics for α -Se₇₅Ge₂₅ positive resist exposed to different doses and incident energies of Ga⁺ ions. Normalized remaining thickness corresponds to $(Z_{min}-\text{etched depth})/Z_{min}$. Wet-chemical development: 1:1:3 HNO₃:HCl:H₂O. Etching time=10 s. Incident energy: -●- 10 keV; -■- 30 keV; -▲- 80 keV.



FIG. 5. SEM photograph of the fabricated sample by a low-energy FIB exposure (6.5×10^{15} ions/cm² at 30 keV) with a multiscan beam diameter of 0.200 μ m and wet-chemical development for 10 s. The linewidth is about 0.225 μ m.

increase simply for increasing FIB energy and both S and γ have maximum values for ~ 70 keV FIBL.

Figure 5 shows a SEM image of resist pattern obtained from low-energy FIB exposure (6.5×10^{15} ions/cm² at 30 keV) with a multiscan diameter of 0.200 μ m. The fabricated linewidth is estimated to be about 0.225 μ m; fabricated depth was about 590 Å.

This linewidth observed is a little larger than the beam diameter (about 250 Å), which is consistent with that of our MC simulation,¹² while in the case of the electron beam the fabricated linewidth is much larger than the beam diameter, e.g., 1.0 μ m linewidth for 0.25- μ m-diam electron beam (8 keV) in Ge-Sb-Se resist.²⁵ This is due to the much smaller lateral spread of ion beam in the resist.

IV. SUMMARY

The ion-induced changes and the exposure response and patterning characteristics of α -Se₇₅Ge₂₅ thin film which can be used as a positive resist in FIBL have been studied using a low-energy Ga⁺ ion source. And then the ion range parameters and the minimum resist thickness are obtained by LSS-based calculation and our MC simulation. This ion resist for a low-energy FIB exposure shows very large ion-induced change in optical absorption properties, which results in an increase of the etch selectivity. In our present work, the mechanism of FIBL depends strongly on the energy density deposited by the collisions of energetic ions rather than the implanted Ga⁺ ions themselves. We, therefore, can expect the improvement of sensitivity and contrast for 70 keV

Ga⁺ FIBL at which the ion-stopping power for the combination of the α -Se₇₅Ge₂₅ resist and Ga⁺ ions exhibits a maximum. Moreover, it can be said that the combination of an α -Se₇₅Ge₂₅ resist and a low-energy FIBL has a great potential for generating patterns down to the sub-0.1- μ m region due to very small lateral spreading effect and that this lithography is applicable to *in situ* processing by utilizing ion-beam-assisted etching or reactive ion etching techniques.

ACKNOWLEDGMENTS

This work was supported by the Inter-University Semiconductor Research Center, Seoul National University, Project No. ISRC-96-E-4036.

- ¹S. Takahashi, M. Ohashi, S. Fukatsu, Y. Shiraki, and R. Ito, *J. Vac. Sci. Technol. B* **11**, 268 (1993).
- ²W. Chu, A. Yen, K. Ismil, M. I. Shepard, H. J. Lezec, C. R. Musil, J. Melngailis, Y.-C. Ku, J. M. Carter, and H. I. Smith, *J. Vac. Sci. Technol. B* **7**, 1583 (1989).
- ³J. Melngailis, *Nucl. Instrum. Methods Phys. Res. B* **80/81**, 1271 (1993).
- ⁴M. A. Gesley, F. J. Hohn, R. G. Viswanathan, and A. D. Wilson, *J. Vac. Sci. Technol. B* **6**, 2014 (1988).
- ⁵J. S. Huh, M. I. Shepard, and J. Melngailis, *J. Vac. Sci. Technol. B* **9**, 173 (1991).
- ⁶K. Balasubramanyam, L. Karapiperis, C. A. Lee, and A. L. Ruoff, *J. Vac. Sci. Technol.* **19**, 18 (1981).
- ⁷E. Nishimura, T. Takigawa, T. Abe, and Y. Katoh, *J. Vac. Sci. Technol.* **4**, 164 (1986).
- ⁸W. M. Moreau, *Semiconductor Lithography: Principle, Practices, and Material* (Plenum, New York, 1988), pp. 409 and 523.
- ⁹J. F. Gibbons, W. S. Johnson, and S. W. Mylroie, *Projected Range Statistics*, 2nd ed. (Dowden, Hutchinson, and Ross, Stroudsburg, PA, 1975).
- ¹⁰H.-Y. Lee and H.-B. Chung, *Synth. Met.* **71**, 2047 (1995).
- ¹¹H.-Y. Lee and H.-B. Chung, *J. Appl. Phys.* **78**, 5975 (1995).
- ¹²H.-Y. Lee and H.-B. Chung, *Jpn. J. Appl. Phys.* **1** **36**, 2409 (1997).
- ¹³J. Lindhard, M. Scharaff, and H. E. Schiott, *K. Dan. Vidensk. Selsk. Mat. Fys. Medd.* **33**, 14 (1963).
- ¹⁴J. Lindhard, V. Nielsen, and M. Scharff, *K. Dan. Vidensk. Selsk. Mat. Fys. Medd.* **10**, 36 (1968).
- ¹⁵N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials*, 2nd ed. (Clarendon, Oxford, 1979), pp. 442 and 547.
- ¹⁶D. L. Dexter and R. S. Knox, *Excitons* (Wiley-Interscience, New York, 1965), p. 321.
- ¹⁷K. Tada, N. Tanio, T. Murai, Y. C. Liang, and K. Furutani, *Thin Solid Films* **108**, 293 (1983).
- ¹⁸Y. Utsugi and Y. Mizushima, *J. Appl. Phys.* **51**, 1773 (1980).
- ¹⁹J. P. deNeufville, *Optical Properties of Solids-New Developments*, edited by B. O. Seraphin (North-Holland, Amsterdam, 1976).
- ²⁰Y. Utsugi and Y. Mizushima, *J. Appl. Phys.* **50**, 1494 (1979).
- ²¹R. Wolfe, *Applied Solid State Science* (Academic, New York, 1975), Vol. 5, p. 165.
- ²²Z. U. Borisova, *Glassy Semiconductors* (Plenum, New York, 1984), pp. 108 and 122.
- ²³P. Andonov, *J. Non-Cryst. Solids* **47**, 297 (1982).
- ²⁴K. L. Chopra and A. Kumar, *Appl. Phys. Lett.* **40**, 5 (1982).
- ²⁵A. Kumar, L. K. Malhotra, and K. L. Chopra, *J. Vac. Sci. Technol. A* **5**, 354 (1987).
- ²⁶N. Koshida, Y. Ichinose, K. Ohtaka, M. Komuro, and N. Atoda, *J. Vac. Sci. Technol. B* **8**, 1093 (1990).

Novel effects in inorganic $\text{As}_{50}\text{Se}_{50}$ photoresists and their application in micro-optics

V. Lyubin,^{a)} M. Klebanov, I. Bar, and S. Rosenwaks

Department of Physics, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

N. P. Eisenberg and M. Manevich

Jerusalem College of Technology, Jerusalem 91160, Israel

(Received 21 August 1996; accepted 9 May 1997)

Some novel properties of one-layer inorganic $\text{As}_{50}\text{Se}_{50}$ photoresists are demonstrated. The photosensitivity of the photoresists increases drastically (1000 times and more) upon exposure to short intense light pulses. The exposure characteristics can be changed from steep to gently sloped ones. The photoresists can be used effectively in both negative and positive modes. Combination of these properties with good optical properties such as a high refractive index, high transparency in the IR and high resolution, makes $\text{As}_{50}\text{Se}_{50}$ photoresists attractive for application in nonconventional optics. Close packed spherical and cylindrical microlens arrays for IR based on the $\text{As}_{50}\text{Se}_{50}$ photoresists are fabricated and tested. The advantage of this technique is in the elimination of thermal reflow and plasma etching that are commonly used. © 1997 American Vacuum Society. [S0734-211X(97)12504-5]

I. INTRODUCTION

Inorganic chalcogenide photoresists (ChPs) have been investigated and applied in photolithography for a long time.¹⁻⁸ Two types of photoresists have been developed: one-layer based on photoinduced structural transformations phenomenon^{9,10} and two-layer consisting of the chalcogenide film and a layer of silver or silver-containing material, based on photoinduced silver diffusion in the chalcogenide glass.^{11,12} In addition to the traditional method of ChPs film preparation by vacuum evaporation,¹⁻⁷ a spin-coating technique was also developed.^{13,14} ChPs are characterized by very high resolution, high contrast, wide spectral range of photosensitivity and the possibility to be used on both planar and nonplanar substrates. These and other interesting properties of such resists make them attractive for application in many photolithographic processes. Wide use of ChPs was limited first of all by their low photosensitivity which was considerably less than the photosensitivity of the modern organic photoresists. Attempts to improve the ChPs by changing their composition, preparation, or irradiation regimes resulted usually in an increase of photosensitivity by not more than two to three times.

This article demonstrates some new properties of the one-layer inorganic $\text{As}_{50}\text{Se}_{50}$ photoresist that open new possibilities in their application. The use of inorganic photoresists in micro-optics for fabrication of microlens arrays for the IR range is also described. The $\text{As}_{50}\text{Se}_{50}$ films were selected for detailed study since a strong photoresist effect was observed in such films previously.¹⁵ Many peculiarities observed in these films are typical also for other ChPs films, for example, for $\text{As}_{40}\text{Se}_{60}$ and $\text{As}_{45}\text{Se}_{45}\text{Te}_{10}$ chalcogenide films.

II. EXPERIMENT

$\text{As}_{50}\text{Se}_{50}$ films were prepared by the evaporation of crushed $\text{As}_{50}\text{Se}_{50}$ bulk glass onto oxide glass or quartz substrates from a quartz crucible in a vacuum of $\sim 5 \times 10^{-6}$ Torr. Film thickness was in the range of 0.1–3.0 μm .

A 250 W Hg lamp, 300 W halogen lamp, 10 mW HeNe laser ($\lambda=633$ nm), excimer ArF laser (Lambda Physik, EMG 101 MSC, $\lambda=193$ nm, single pulses of 0.5–4.5 mJ, pulse width $t=16$ ns), and Nd:YAG laser ($\lambda=266$ nm, $t=5$ ns) were used as sources of irradiation. The rate of dissolution was determined by monitoring the local thickness of the film using thin film interference.¹⁶ In these experiments solutions based on ethylenediamine and ethanolamine were used as negative selective developers. KOH solutions in water or in alcohol were used as positive developers.

III. PHOTORESISTS' PROPERTIES AND POSSIBLE APPLICATION

A. Short intense pulse exposure

It was shown in this work that using short intense pulse exposures leads to a large increase of ChPs photosensitivity. This conclusion was made after direct measurements of the dissolution rates of exposed and unexposed areas of $\text{As}_{50}\text{Se}_{50}$ films.

Figure 1 shows typical characteristics of the dissolution in negative solvent of the $\sim 0.4 \mu\text{m}$ $\text{As}_{50}\text{Se}_{50}$ film areas exposed to cw HeNe laser beam ($E=235$ J/cm²) and to single light pulse of the ArF laser ($E=24$ mJ/cm²). The initial rates of dissolution of exposed areas are considerably lower than those of unexposed parts of the film. Both types of exposure gave approximately the same effects, although the exposure doses in the two cases were substantially different. In other experiments, 10^3 – 10^4 times increase in photosensitivity of studied photoresists was observed when excimer laser pulses

^{a)}Electronic mail: vlubin@bgumail.bgu.ac.il

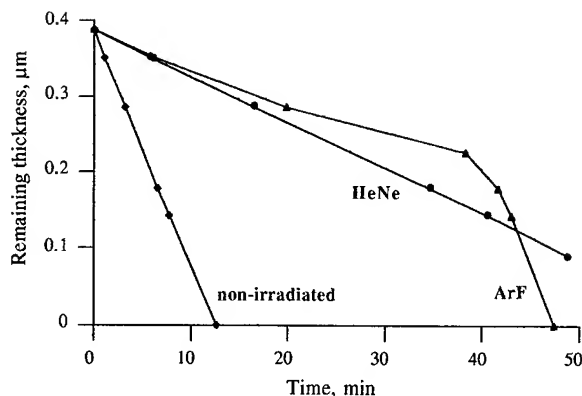


FIG. 1. Thickness of $\text{As}_{50}\text{Se}_{50}$ film areas (irradiated and nonirradiated) vs time of dissolution in ethanolamine.

were used. Films exposed to HeNe laser with a dose of $20\text{--}30\text{ mJ/cm}^2$, equal to the dose obtained by a single pulse of the ArF laser, show the dissolution rate of exposed area practically the same as that of unexposed part of the film.

It is seen from Fig. 1 that after 40 min of dissolution, the dissolution rate of the ArF pulse exposed area increases and becomes comparable to the dissolution rate of unexposed area. This fact is due to the large absorption coefficient for ultraviolet radiation in As-Se films ($\alpha > 5 \times 10^5\text{ cm}^{-1}$ as compared with $\alpha \sim 2 \times 10^4\text{ cm}^{-1}$ at $\lambda = 633\text{ nm}$) so that the excimer laser pulse affects the film only to a depth of about $0.15\text{ }\mu\text{m}$. Therefore for the most effective use of such resists with an ArF irradiation source, the resist thickness must not exceed $0.15\text{ }\mu\text{m}$. It is interesting to note the following peculiarities:

- (1) The dissolution rate of exposed areas was essentially unchanged when more than one pulse of ArF laser light was used.
- (2) As opposed to negative developer use, the initial rates of dissolution of exposed and unexposed areas differed very slightly when a positive developer was used.

A direct comparison of photosensitivity of $\text{As}_{50}\text{Se}_{50}$ photoresist and Shipley AZ-1400-31 resist under 5 ns pulses at 266 nm (the fourth harmonics of the Nd:YAG laser) was carried out. The chalcogenide photoresist was able to record a single pulse. However, the Shipley photoresist displayed a very weak image (after development), but only after 100 pulses. Thus, the inorganic ChPs, which are substantially less sensitive than Shipley resists at cw radiation, proved to be much more sensitive to high intensity short pulse radiation.

The main result is the demonstration of very large increase ($10^3\text{--}10^4$ times) in the inorganic $\text{As}_{50}\text{Se}_{50}$ photoresists' photosensitivity. This increase can possibly be attributed to the exposure to UV light quanta of large energy (4.7–6.4 eV of the ArF layer versus 2.0 eV of the HeNe laser). This energy difference can drastically change the mechanism of photostructural transformations, that are the basis of photostimulated change of dissolution rate. For example, the high energy quanta could excite the bonding electrons of selenium rather than their lone pair electrons,¹⁷ increasing

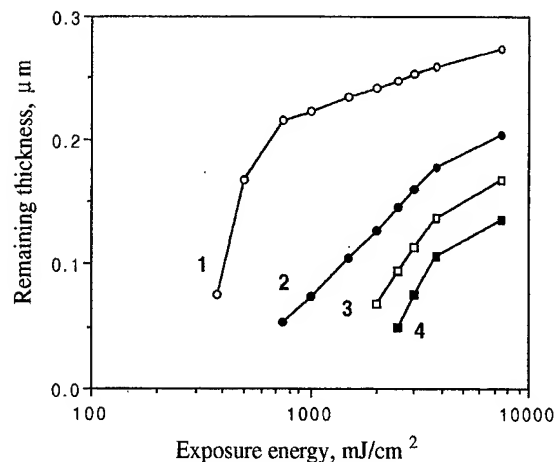


FIG. 2. Contrast characteristics of $0.3\text{ }\mu\text{m}$ $\text{As}_{50}\text{Se}_{50}$ negative photoresist exposed to 250 W Hg lamp light. Development times: (1) 180 s; (2) 400 s; (3) 500 s; and (4) 580 s. Development time for complete dissolution of the unexposed areas: 45 s.

the probability of interatomic bonds breaking and accelerating the process of photostructural transformations. Another possible explanation of the photosensitivity increase, is the appearance of nonlinear recombination and relaxation phenomena under intense short pulse light excitation. Such nonlinear phenomena are displayed also in the nonlinear photoconductivity of chalcogenide glassy semiconductors.¹⁸ These phenomena are due to redistribution of an enlarged amount of the excited charge carriers into different trapping and recombination states that are characteristic of the chalcogenide glasses.¹⁸ Such redistribution leads to the increase of charge carriers lifetime and to the growth of photostructural transformations chances.

B. Exposure characteristics

The exposure characteristic of a photoresist, indicating the remaining photoresist thickness as a function of exposure energy, is very important for any photolithographic application. For most organic photoresists a steep exposure curve is typical, and this steepness is particularly suitable for the application of photoresists in microelectronics. The form of this characteristic can be only slightly changed by adjustment of exposure and development conditions.

On the contrary, for inorganic ChPs, the exposure characteristics are slightly sloped and their form can be changed significantly when the exposure conditions (light spectrum, maximum light intensity, exposure time) and development conditions (kind of the developer, development time) are varied. Several exposure characteristics of an $\text{As}_{50}\text{Se}_{50}$ photoresist are shown in Fig. 2 and indicate the possibility of realizing half-tones when the remaining resist thickness is different in the areas exposed by light of varied intensity. This behavior of inorganic photoresists is due to the property of photostructural transformations to be accumulated in a chalcogenide film during exposure.

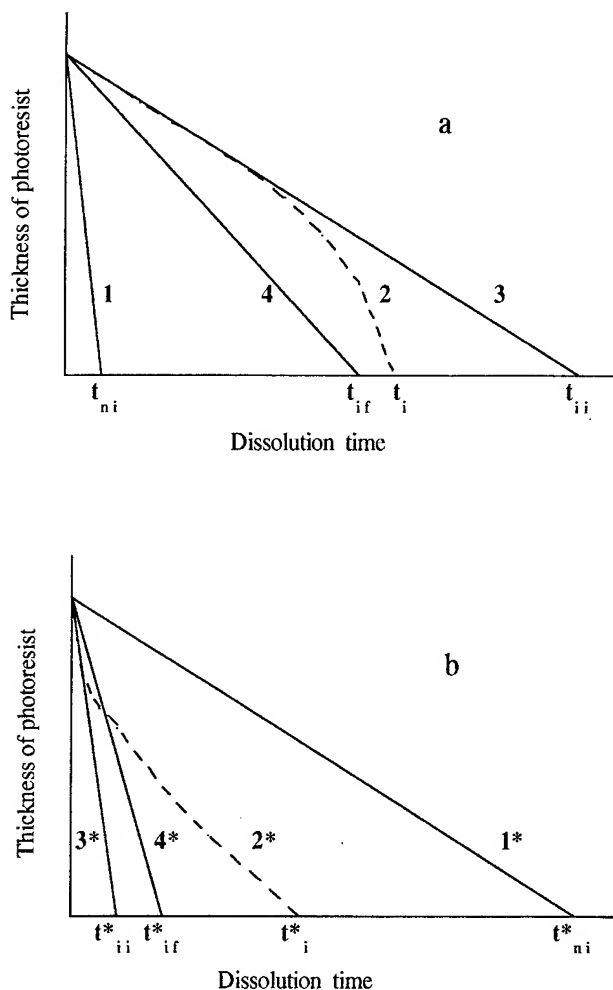


FIG. 3. Schematic representation of dissolution kinetics of nonirradiated (1,1*), irradiated by polychromatic light (2,2*), and irradiated through band-pass filter (4,4*) areas of the photoresist for (a) negative and (b) positive developers. Lines 3,3*, related to the seeming case of nonabsorbing photoresist, are obtained by continuation of initial parts of curves 2 and 2*.

C. Positive and negative modes of chalcogenide photoresist

The possibility of the easy transition of a one-layer ChPs from positive to negative mode was shown previously.^{7,15} Such a transition is accomplished by a simple change of the developer composition and is based on the specific polymer structure of the chalcogenide glasses in which the irradiation induces, simultaneously, both photodestruction and photopolymerization processes. During these processes, some interatomic bonds are partially broken and other bonds are created. In this case, the developer interacting with broken bonds acts as a positive one, while the developer interacting with created bonds leads to the negative effect.⁷

At the same time, an essential difference has been observed by many investigators between positive and negative modes in most of the ChPs. For the negative mode large values (up to 20–30) of the γ factor, that is the ratio of dissolution rates in unexposed and exposed areas, were achieved for resists of As–Se glassy system, while for the

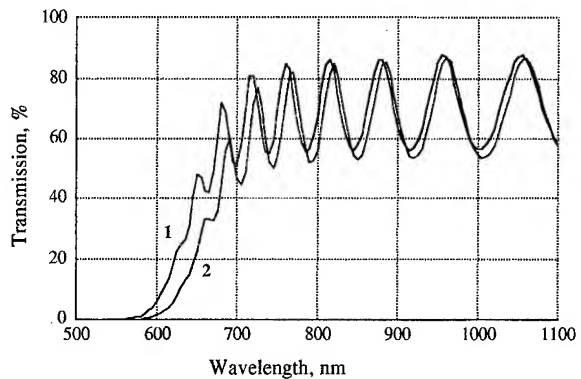


FIG. 4. Transmission spectra of (1) nonirradiated and (2) irradiated $\sim 2.5 \mu\text{m}$ $\text{As}_{50}\text{Se}_{50}$ film.

positive mode, only small values (less than three) of the γ factor (ratio of dissolution rates in exposed and unexposed areas in this case) are reported.¹⁵ This peculiarity can be understood if the property of ChPs to decrease the transpar-

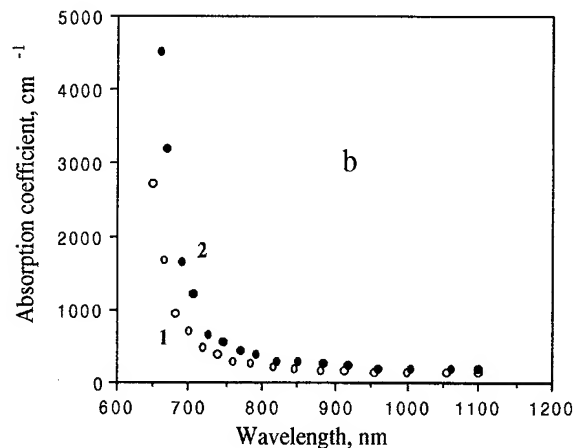
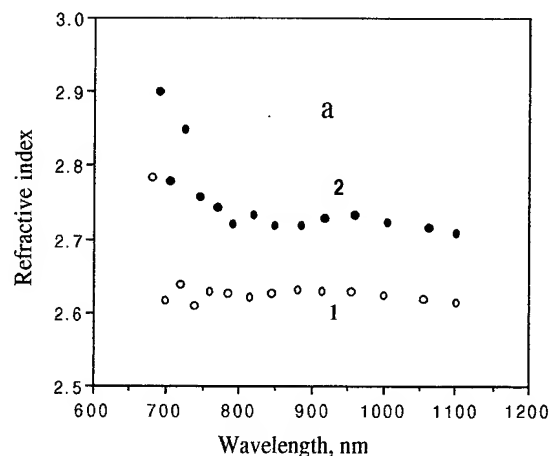


FIG. 5. Spectra of (a) refractive index and (b) absorption coefficient of (1) nonirradiated and (2) irradiated $\sim 2.5 \mu\text{m}$ $\text{As}_{50}\text{Se}_{50}$ film.

ency under light irradiation, the photodarkening effect,^{9,10} is taken into consideration.

Due to photodarkening, the depth of exciting light penetration into the volume of photoresist film is diminished, and this effect leads to different results in the positive and negative modes. This situation is explained schematically in Fig. 3 where the kinetics of nonirradiated and irradiated photoresist film dissolution is shown. Here the t_{ni} and t_{ni}^* , t_i and t_i^* , t_{ii} and t_{ii}^* are the times of total dissolution of nonirradiated, irradiated real, and irradiated ideal (nonabsorbing) areas of negative and positive (*) photoresist, respectively. The γ factor values for real and ideal (nonabsorbing) cases can be written for the negative mode as $\gamma = t_i/t_{ni}$ and $\gamma_i = t_{ii}/t_{ni}$, and for the positive mode as $\gamma^* = t_{ni}^*/t_i^*$ and $\gamma_i^* = t_{ni}^*/t_{ii}^*$. It is seen from the figure that in the ideal case, both γ_i and γ_i^* can be large enough, but in the real case of absorbing ChPs, $\gamma^* \ll \gamma$, in good agreement with the situation observed experimentally. So, in the case of As₅₀Se₅₀ resists, values $\gamma^* = 2-3$ were recorded.

It was shown that the situation for the positive mode can be improved if the strongly absorbed short wavelength light is filtered. Curves 4 and 4* in Figs. 3(a) and 3(b) indicate schematically the kinetics of dissolution of photoresist film irradiated from the same light source through a bandpass filter, cutting off the short wavelength radiation. Corresponding, γ factor values can be written as $\gamma_f = t_{if}/t_{ni}$ and $\gamma_f^* = t_{ni}^*/t_{if}^*$. In the negative mode γ_f is slightly smaller than γ , but in the positive mode the γ_f^* value is substantially larger than γ^* . Experiments with As₅₀Se₅₀ photoresist using a halogen lamp as the irradiation source and application of KS-10 bandpass filter (transmits light with $\lambda > 0.6 \mu\text{m}$) led to an approximate three times increase of the γ factor, that achieved values of six through eight.

D. Optical properties

Optical properties of ChPs are substantially different from those of organic photoresists. The ChPs are characterized by high values of refractive index $n=2.2-2.9$ and by good transparency in the IR spectrum range.

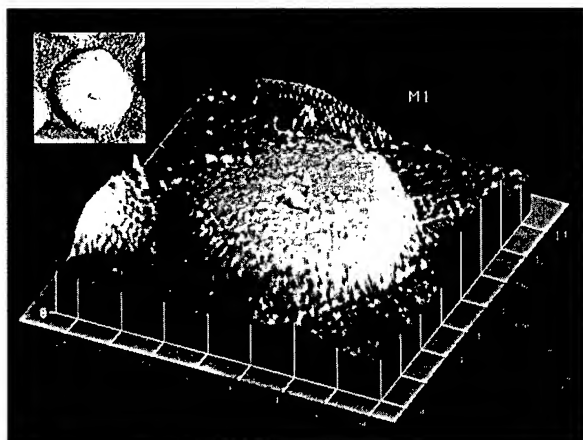


FIG. 6. Atomic force microscope image of one spherical microlens.

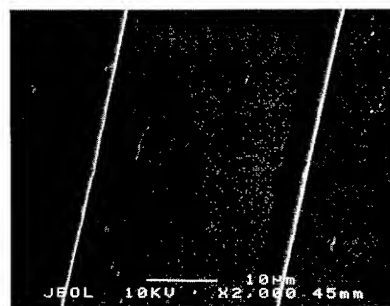


FIG. 7. Scanning electron microscope micrograph of section of two long cylindrical microlenses.

The transmission spectrum of a typical As₅₀Se₅₀ photoresist film is shown in Fig. 4 for the visible and near IR range. The positions of the interference extrema in Fig. 4 depend on resist thickness. The values of refractive indices and absorption coefficients calculated from the transmission spectrum using the Swanepoel method¹⁹ are plotted in Fig. 5. The photodarkening effect is displayed in the change of transmission spectrum and in the increase of refractive index and absorption coefficient. All these changes for As₅₀Se₅₀ photoresist film are illustrated in Figs. 4 and 5.

E. Prospects of photoresists application

The above-mentioned peculiarities of As₅₀Se₅₀ photoresist differentiate them from traditional organic photoresists. Exposure by short intense light pulses, accompanied by a strong increase of photosensitivity of ChPs, removes their main drawback and makes them promising for application in sub-micron lithography and in nanolithography, particularly, in near-field lithography²⁰ where short light pulses are used. As₅₀Se₅₀ photoresists that are characterized by high sensitivity to short UV pulses of ArF excimer laser seem to be especially interesting in connection with the recent interest of modern microelectronics in 193 nm lithography.²¹

The ability to change the form of exposure characteristics of As₅₀Se₅₀ photoresists allows their use for the fabrication of half-tone photomasks and for various applications in non-conventional optics. The possible use of As₅₀Se₅₀ photoresists both in the negative and positive modes shown here is also useful for different applications.

The As₅₀Se₅₀ films possess simultaneously interesting photoresist properties and good optical properties, such as a high refractive index and high transparency in the IR. Due to this combination, such films can be used in nonconventional optics, particularly, for fabrication of microlens arrays and diffractive optical elements.

IV. APPLICATION FOR FABRICATION OF IR MICROLENS ARRAYS

The microlens arrays were fabricated mainly in As₅₀Se₅₀ films that were deposited onto glass, silicon, and gallium arsenide substrates and had thicknesses from 0.3 to 1.0 μm . Proximity contact printing lithography²² was used. Planoconvex spherical IR microlens arrays were fabricated on silicate glass substrates. A He-Ne laser was used as a light source.

This radiation was approximately homogeneously absorbed in the As₅₀Se₅₀ film, ensuring high contrast in dissolution of irradiated and nonirradiated areas of the film. The development of exposed photoresist was made in the negative photoresist mode using the ethylenediamine-based developers. The microlens arrays, created in the chalcogenide photoresist, were obtained in their final form without the need of transfer into the substrate material by standard plasma etching that is usually used in the case of organic photoresists application.²³

An array produced with the 0.32 μm thick resist gave the following results:

Refractive index ($\lambda=0.8 \mu\text{m}$): 2.9	Pitch: 12 μm
Diameter of single lens: 12 μm	Array size: 40×40 microlenses
Focal length ($\lambda=0.5 \mu\text{m}$): 30 μm	

The exposure doses of 5–10 J/cm² and dissolution times of 3–5 min in the negative developer (8% amine-based solution) were used in the fabrication of such microlens arrays. An atomic force microscope image of one microlens is shown in Fig. 6.

The experimental focal length value f was very close to the value determined from the known formula for a plano-convex lens $f=r^2+h^2/2h(n-1)$, where r is the radius of the lens, n is the refractive index, and h is the maximum thickness of chalcogenide film.

Long cylindrical lenses (few millimeters) were also fabricated using the same lithographic technique with As₅₀Se₅₀ films of thickness 0.45 μm . In Fig. 7 are shown cylindrical lenses with the following parameters:

Refractive index ($\lambda=0.8 \mu\text{m}$): 2.9;
Lens width: 4.2 μm ; and
Focal length ($\lambda=0.5 \mu\text{m}$): 3.0 μm .

V. CONCLUSION

The detailed study of As₅₀Se₅₀ photoresist films prepared by resistive heating vacuum evaporation resulted in the revelation of several new features. Large photosensitivity increase when short intense light pulse exposure is used, the

ability to change the exposure characteristics, and use as a positive photoresist were shown. Most of these features are explained on the base of general knowledge of the photoinduced phenomena in chalcogenide glasses. These peculiarities together with their interesting optical properties, make this type of inorganic photoresists attractive for different applications.

Fabrication of close-packed microlens arrays for the IR, using As₅₀Se₅₀ photoresists has been shown. In the method used, thermal reflow and ionic etching were eliminated, making this fabrication process very attractive.

ACKNOWLEDGMENT

This work was supported by a grant from the Israeli Ministry of Science and the Arts.

- ¹H. Nagai, A. Yoshikawa, Y. Toyoshima, O. Ochi, and Y. Mizushima, *Appl. Phys. Lett.* **28**, 145 (1976).
- ²M. S. Chang and J. T. Chen, *Appl. Phys. Lett.* **33**, 892 (1978).
- ³K. L. Tai, W. R. Sinclair, R. G. Vadimsky, J. M. Moran, and M. J. Rand, *J. Vac. Sci. Technol.* **16**, 1977 (1979).
- ⁴B. Singh, S. Rajagopalan, and K. L. Chopra, *J. Appl. Phys.* **51**, 1768 (1980).
- ⁵A. G. Poleshuk, E. G. Churin, Y. I. Yurlov, and B. Mednikarov, *J. Imaging Sci.* **30**, 132 (1986).
- ⁶G. H. Bernstein, W. P. Liu, Y. N. Khawaja, M. N. Kozicki, D. K. Ferry, and L. Blum, *J. Vac. Sci. Technol. B* **6**, 2298 (1988).
- ⁷B. T. Kolomiets, V. M. Lyubin, and V. P. Shilo, *Phys. Chem. Glasses* **4**, 351 (1978).
- ⁸S. A. Dumford and J. M. Lavin, *J. Vac. Sci. Technol. B* **12**, 44 (1994).
- ⁹S. R. Elliott, *J. Non-Cryst. Solids* **81**, 71 (1986).
- ¹⁰V. M. Lyubin, *J. Non-Cryst. Solids* **97&98**, 47 (1987).
- ¹¹A. V. Kolobov and S. R. Elliott, *Adv. Phys.* **40**, 625 (1991).
- ¹²K. Tanaka, *J. Non-Cryst. Solids* **137&138**, 1021 (1991).
- ¹³G. C. Chern and I. Lauks, *J. Appl. Phys.* **53**, 6979 (1982).
- ¹⁴S. Shtutina, M. Klebanov, V. Lyubin, S. Rosenwaks, and V. Volterra, *Thin Solid Films* **261**, 263 (1995).
- ¹⁵V. M. Lyubin, A. M. Sedikh, N. N. Smirnova, and V. P. Shilo, *Microelectronica* **18**, 523 (1989).
- ¹⁶M. S. Chang and T. W. Hou, *Thin Solid Films* **55**, 463 (1978).
- ¹⁷M. Kastner and H. Fritzsche, *Philos. Mag. B* **37**, 199 (1978).
- ¹⁸B. T. Kolomiets and V. M. Lyubin, *Phys. Status Solidi A* **17**, 11 (1973).
- ¹⁹R. Swanepoel, *J. Phys. E* **16**, 1214 (1983).
- ²⁰K. Lieberman and A. Lewis, *Appl. Phys. Lett.* **62**, 1335 (1993).
- ²¹T. Moss, *Photonic Spectra* **30**, 27 (1996).
- ²²W. M. Moreau, *Semiconductor Lithography* (Plenum, New York, 1988).
- ²³E. J. Gratz, *Proc. SPIE* **1992**, 266 (1993).

Imaging characteristics of poly(methyl methacrylate) at vacuum ultraviolet wavelengths

I. E. Ferincz,^{a)} Cs. Tóth,^{b)} and J. F. Young^{c)}

Department of Electrical and Computer Engineering and the Rice Quantum Institute, Rice University, Houston, Texas 77251

(Received 11 October 1996; accepted 9 May 1997)

We report the depth versus exposure characteristics of poly(methyl methacrylate) (PMMA) in the vacuum ultraviolet (VUV) region from 59 to 128 nm. Calculated absorption coefficients of the PMMA are also presented. The depth of the features on the PMMA were in the range of 20–80 nm created by 1–50 mJ cm⁻² exposures and were measured with an atomic force microscope (AFM). We found that the AFM can reliably measure depth differences of only 2 nm. The sensitivity of the AFM permits exposures to be reduced for high resolution microscopy and holography even in the case of highly absorbing films. © 1997 American Vacuum Society. [S0734-211X(97)12404-0]

I. INTRODUCTION

Microscopy using contact printing or holography requires a well characterized, high resolution recording film. One of the earliest and best characterized electron and x-ray resists is poly(methyl methacrylate) (PMMA).^{1–3} In contrast, however, to lithography applications — where the full removal of the resist is desirable in the exposed areas — contact imaging and holography require a resist with vertical (depth) resolution. The attainability of a “grey-scale” pattern in the resist is important in these applications for reliable and reproducible mapping of the light intensity in a wide exposure range. In the soft x-ray region, microscopy using PMMA as a resist material is well developed.⁴ In PMMA, a positive resist, bonds are broken by incident short wavelength radiation, decreasing the average molecular weight of the long-chain molecules. This increases the solubility of the resist in a suitable solvent, resulting in a pattern of surface heights corresponding to the exposure. The surface pattern can be measured by a scanning⁵ or transmission electron microscope,^{6,7} but an atomic force microscope (AFM) provides more reliable quantitative data.^{8,9}

Vacuum ultraviolet (VUV) microscopy has the potential to supplement x-ray and electron microscopy, because the contrast mechanism is quite different. VUV radiation interacts with the outer valence electrons and molecular absorption is sensitive to composition and structure. In addition, several reasonably practical incoherent and coherent sources exist in the 80–130 nm region. However, almost no studies have been made of PMMA, or other films, for VUV imaging. While there are lithography studies in the 110–260 nm region,^{10–12} this information is only partly relevant since the requirements of lithography are quite different from imaging. For lithography, the critical parameter is the exposure required to completely remove the film; imaging, however, requires quantitative knowledge of the response function and

a reasonable dynamic range, or intensity level resolution. The relationship between exposure and surface height is generally poorly known even for common photoresists. Basic and important optical constants, such as absorption coefficient at VUV wavelengths are also lacking.

In this work we will present the depth-exposure characteristics of PMMA in the 59–128 nm spectral range, for exposure values of 1–50 mJ cm⁻². The surface depths were measured with an AFM and were in the range of 2–80 nm. Since the AFM can reliably measure surface height differences of only 2 nm, very low exposures can be used even at wavelengths that are highly absorbing. We have used our depth versus exposure data to calculate the VUV absorption coefficients of PMMA using a simple model for exposure and dissolution.

II. EXPERIMENTAL DESCRIPTION

A. Preparation and development of the PMMA resist

The samples for exposure were 12.5 mm diameter polished glass disks spin coated with PMMA. The resist was a commercial solution of 4% PMMA (molecular weight of 950 000) in chlorobenzene.¹³ The top surface of the substrate was flooded with the PMMA solution and then spun at 5000 rpm for 25 s. A second layer of PMMA was spun over the first to provide a thicker film. The resulting film was 290 ± 10 nm thick as measured by the AFM. The substrate and film was then baked at 140 °C for at least 30 min in a vacuum oven at ~500 mbar pressure.

Following exposure the film was developed in pure methyl isobutyl ketone for 60 s at 25 ± 0.5 °C, during which time the developer was agitated mildly. Following development, the sample was immediately rinsed in isopropyl alcohol for 5 s, dried with a jet of clean freon, and baked again at 140 °C for 30 min.

B. System setup

Figure 1 shows the experimental arrangement. Broadband incoherent VUV light was generated using a laser-produced plasma. Pulses from a Q-switched Nd:YAG laser were fo-

^{a)}Present address: Department of Optics and Quantum Electronics, JATE University, H-6720 Szeged, Dóm tér 9, Hungary.

^{b)}Permanent address: Research Institute for Solid State Physics, P.O. Box 49, H-1525, Budapest, Hungary.

^{c)}Electronic mail: young@rice.edu

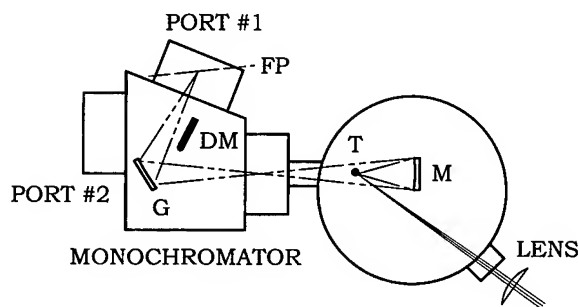


FIG. 1. Experimental arrangement for the resist characterization. The incoming Nd:YAG laser pulse (solid lines) was focused to the target rod T, and the VUV radiation (dash-dot-dot lines) was collected and imaged by the gold mirror M. The grating G filtered and focused the VUV light to the focal plane FP, which was tangential to the Rowland circle. The energy could be monitored during the exposure by turning the diverter mirror DM into and out from the beam path.

cused onto a 6 mm diameter stainless steel rod with a 30 cm focal length lens. The 930 ± 10 mJ and 7 ns full width at half maximum (FWHM) pulses produced infrared light intensity of 10^{12} W cm $^{-2}$ at the target surface. The target rod was attached to and rotated by the motorized screw¹⁴ to present a fresh target surface for each laser pulse. The laser repetition rate was 10 Hz and this arrangement allowed 75 min long exposures. The light from the laser-produced plasma was collected by a 2 in. diameter spherical gold mirror (focal length of 5 cm) and imaged on the input slit of a 0.2 m, f #4.5 monochromator. High purity helium at 20 Torr was flowed through the chamber during the exposures to reduce the accumulation of plasma debris on the mirror.

For the PMMA exposures, the monochromator was modified by replacing the output slit at Port No. 1 with a sample holder tangent to the Rowland circle, allowing us to make exposures at several wavelengths simultaneously. The monochromator grating had 1200 lines/mm leading to a plate factor of 3.4 nm/mm in the focal plane. Exposures were made through a mask (see Fig. 2) with identical patterns spaced 1.25 mm apart, corresponding to 4.25 nm in wavelength. The five patterns spanned a total wavelength region of 17 nm. The input slit was set at 1.25 mm for all exposures. The mask was fabricated from 75 μ m thick polyimide sheet by ablation using a KrF excimer laser. As seen in Fig. 2, each of

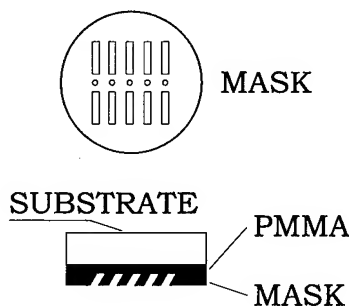


FIG. 2. Front view of the polyimide mask and cross section of the substrate, film, and mask assembly.

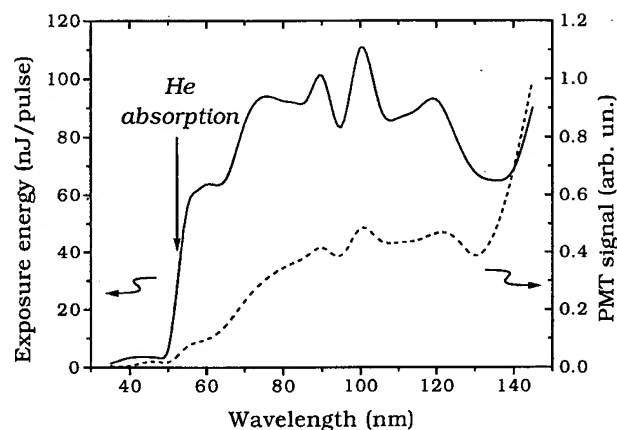


FIG. 3. Typical spectrum of the laser produced plasma light source used in the experiments, as measured directly by the PMT at Port No. 2 (dashed line), and the calibrated light energy deposited on the sample surface at Port No. 1 (solid line).

the five mask patterns consist of two bars, 60 μ m wide by 2.5 mm long, separated by a 20 μ m diameter circle. Film depth measurements were made on the circular dot exposures; the bars served only as visual aids to locate the circles. The mask was placed directly against the PMMA film and the assembly held in the monochromator's focal plane.

C. Calibration of VUV exposure flux

The determination of the depth versus exposure characteristics of PMMA requires careful measurement of the exposure flux over a range of wavelengths. Our primary energy standard was the well studied and stable quantum efficiency of Al $_2$ O $_3$ cathode photodiodes.¹⁵ Because the sensitivity of this photocathode drops sharply for wavelength longer than 100 nm, we used it primarily to calibrate a sodium salicylate-photomultiplier combination (PMT) that was used to measure the source spectrum and to monitor exposure values. The variation of the quantum efficiency of sodium salicylate is about 5% over our 35–130 nm wavelength measurement range.¹⁵

As a first step of calibration, we measured the energy at monochromator Port No. 1 with the Al $_2$ O $_3$ photodiode in the 60–100 nm wavelength range. During the calibration the input and output slits were opened to 100 μ m. The spectrum was then recorded by the PMT at the same port and calibrated by the results from the Al $_2$ O $_3$ -photodiode measurement. Next, the spectrum was recorded at Port No. 2 and the reflection of the diverter mirror was calibrated by comparing the two spectra. Using this method we were able to evaluate the energy at Port No. 1 from the measured signals of the PMT at Port No. 2 at each wavelength between 50 and 130 nm (Fig. 3). The exposure energies varied between 60 and 120 nJ per laser pulse. We estimate that the error of the absolute energy values is 20%, while the relative error between different wavelengths is only 12%. The energy density at the sample plane was between 7.0×10^{-4} and 1.4×10^{-3} mJ cm $^{-2}$ per pulse, as calculated by dividing the measured energy by the effective open area of the output slit. The total

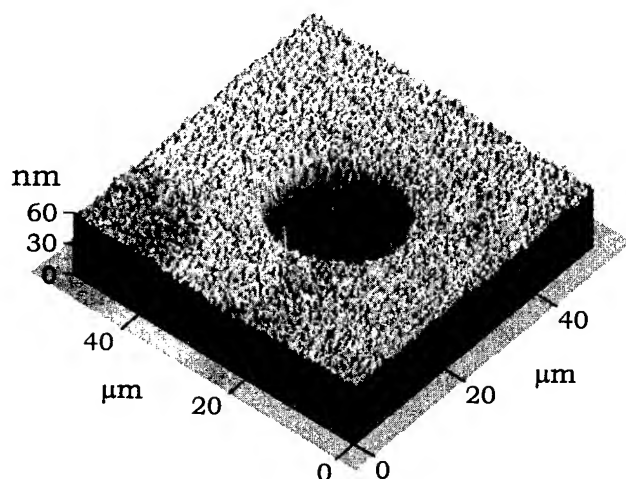


FIG. 4. An atomic force micrograph of a contact printed pattern on the surface of PMMA.

exposure was obtained by simply multiplying this average exposure value per laser pulse by the total number of pulses in an exposure.

In order to account for possible long-term changes in the efficiency of the laser-produced plasma and the light collection system, we interrupted each exposure to take source spectra at Port No. 2. The number of spectra taken was varied to achieve good average values. Three spectra were measured for the shortest (3 and 4 min) exposures and seven for the longest (60 min) exposures. The exposure energy was calculated by averaging the spectra.

D. Depth measurement by AFM

The transverse and vertical scans of the AFM¹⁶ were calibrated using a calibration grating and height standards. A typical atomic force micrograph of a contact printed test pattern on the surface of the PMMA is shown in Fig. 4. To obtain a good estimate of the depth of a given hole, four $\sim 50 \times 50 \mu\text{m}$ (256 by 256 pixels) scans were made using different scanning directions and the results averaged. Typically, the standard deviation of the pixel height from the average was $\sim 3 \text{ nm}$. However, because of the large number of pixels, this deviation did not effect the depth measurement directly, since the *average* height of the pixels were compared. However, this high frequency noise did introduce errors during the removal of the apparent surface curvature. Apparent surface curvature is an inherent byproduct of the AFM transverse scanning mechanism, and it can be well approximated by a second order polynomial. The elimination of the curvature was performed by fitting a second order polynomial to each data row of the scans. The accuracy of this fit is reduced by the pixel height variation because of the small number of pixels in each row. This was particularly true where the data row contained values from both the exposed and unexposed parts, since the unexposed part was not included in the fitting. We estimate that the vertical resolution was limited to 2 nm by this effect.

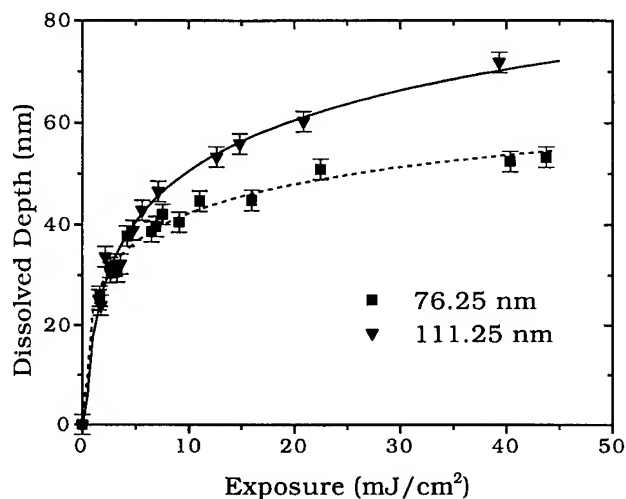


FIG. 5. Developed film depth vs exposure for PMMA at two different wavelengths. The symbols with error bars are the measured height differences between exposed and unexposed regions, while the lines are the results of a model fitted to the data.

III. RESULTS AND DISCUSSION

Measurement of exposure patterns such as Fig. 4 resulted in characteristic curves of depth versus exposure, S , at specific wavelengths spaced 4.25 nm over the range of 59–128 nm. Figure 5 shows two such curves; the symbols represent measured data while the lines are the result of the exposure-dissolution model discussed below. This set of curves can be usefully summarized by defining a specific exposure flux, S_p , for each wavelength that will produce a 30 nm deep pattern. Figure 6 shows a plot of S_p over the wavelength range measured. While the 30 nm criteria is arbitrary, we found that patterns of this depth resulted in low noise, reproducible measurements. Also shown in Fig. 6 is the calculated derivative of pattern depth versus exposure, $G = dz/dS$, evaluated at S_p . The parameter G is a measure of the exposure resolution; a large G indicates there are more distinguishable intensity levels within a given exposure range.

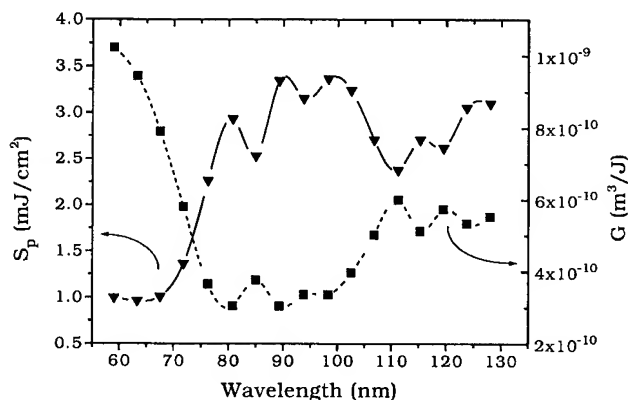


FIG. 6. Specific exposure S_p required for a developed depth of 30 nm, (triangles), as a function of wavelength, and G , the derivative of depth with respect to exposure at S_p (filled squares), characterizing the possible grey-scale levels for imaging applications.

TABLE I. Values of α and κ determined by fitting the model of Eq. (3) to the measured dissolved pattern depth. The wavelength independent value of $n=2.43$.

λ_c (nm)	α (μm^{-1})	κ ($\text{m}^{2n+1}/\text{J}^n\text{s}$)
59.00	98	5.3×10^{-10}
63.25	110	1.3×10^{-9}
67.50	126	3.5×10^{-9}
71.75	126	1.7×10^{-9}
76.25	121	3.6×10^{-10}
80.75	112	9.8×10^{-11}
85.00	105	8.8×10^{-11}
89.25	98	2.8×10^{-11}
93.75	94	2.5×10^{-11}
98.25	88	1.5×10^{-11}
102.50	78	7.8×10^{-12}
106.75	73	9.2×10^{-12}
111.25	70	1.0×10^{-11}
115.25	72	8.4×10^{-12}
119.50	66	6.3×10^{-12}
123.75	61	3.1×10^{-12}
128.00	57	2.4×10^{-12}

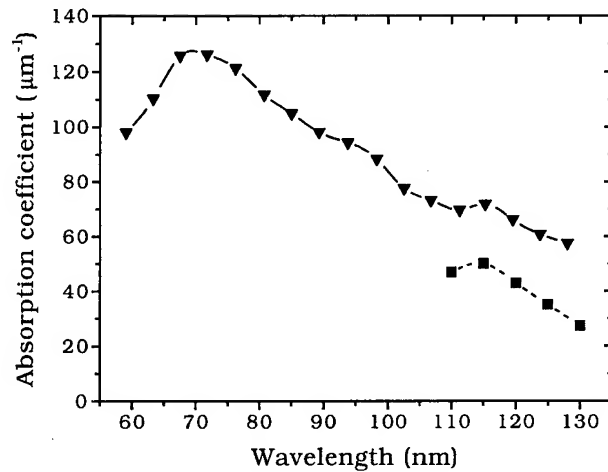


FIG. 7. Absorption coefficient of PMMA as a function of the wavelength (triangles) calculated by fitting the measured pattern depths to the model of Eq. (3). The data shown by filled squares were determined by Kudo *et al.*, by direct absorption measurement (Refs. 10 and 11).

A. Dissolution model of PMMA

We used a simple model of exposure energy deposition and the resulting dissolution of PMMA to calculate the VUV absorption coefficient of PMMA from our pattern depth data. First, we assumed a simple exponential absorption in PMMA so that the energy deposited at any depth z is $D(z) = \alpha S \exp(-\alpha z)$, where α is the absorption coefficient, and S is the incident energy flux at the surface $z=0$. Next, we used the previously proposed model of polymer dissolution^{2,3} that postulates the dissolution rate is proportional to a power of the energy deposited per unit volume:

$$\frac{dz}{dt} = [kD(z)]^n + R_0, \quad (1)$$

where n is a constant and R_0 is the dissolution rate of the unexposed polymer. The parameter k represents the bond breaking efficiency of the radiation and is expected to be a function of wavelength.³ Substituting for $D(z)$, the change of surface height with development time is

$$\frac{dz}{dt} = \kappa S^n e^{-\alpha z(t)n} + R_0, \quad (2)$$

where $\kappa = \alpha^n k^n$. Assuming $z(0)=0$ and $t=0$ at the start of the development, the solution of Eq. (2) is

$$z(t) = \frac{1}{n\alpha} \ln \frac{R_0 e^{\alpha R_0 t n} + (e^{\alpha R_0 t n} - 1) \kappa S^n}{R_0}. \quad (3)$$

This model neglects temperature dependence of the dissolution rate and any dependence of the exponent n on deposited energy.^{2,17}

B. Resist characteristics

We evaluated R_0 by immersing unexposed PMMA samples in methyl isobutyl ketone for various times and measuring the change in thickness with the AFM. We found

that $R_0 = 72.3 \pm 18 \text{ pm s}^{-1}$, which agrees well with the value of 80 pm s^{-1} of Hawryluk *et al.*² Equation (3) was fitted to the experimental data in order to determine the parameters $\alpha(\lambda)$, $\kappa(\lambda)$, and n . The exponent n should depend only on the polymer and solvent properties and it was assumed to be independent of wavelength in the fitting process. Therefore, a simultaneous fitting has been performed for the data from the 17 different wavelengths in order to find the best $(17 \times 2) + 1$ parameters (α and κ for each λ , and one common n) for all the experimental data. Using this method of simultaneous fitting of all the wavelength data — instead of a series of independent three-parameter fittings for each λ — we were able to fit our developed depth versus exposure and wavelength data very well; two examples are shown in Fig. 5. Since the data values in Fig. 5 represent the difference in height of exposed and unexposed areas, the curves plotted are $z(t)$ from Eq. (3) minus $R_0 t$, the amount dissolved from the unexposed regions. For our conditions the correction $R_0 t = 4.3 \text{ nm}$. Values of $\alpha(\lambda)$ and $\kappa(\lambda)$ are given in Table I for each wavelength measured, and the absorption coefficient is plotted in Fig. 7. The 95% confidence interval of the fitted parameters have also been determined: they are $\pm 4\%$, $\pm 11\%$, and $\pm 31\%$, for the values of n , α , and κ , respectively. The optimum value of n was found to be 2.43. Previous studies using x-ray exposure found values between 3 and 3.4, while electron beam exposure yielded values of 1.33 to 2.58.² The calculated absorption coefficient in Fig. 7 is systematically higher than the values measured by Kudo *et al.*,^{10,11} in the range 110–130 nm. This discrepancy might indicate the need for a more comprehensive dissolution model, e.g., incorporation of the D dependence of n .

IV. SUMMARY AND CONCLUSIONS

We have presented the depth versus exposure characteristics of PMMA over a broad VUV spectral region, 59–128 nm. These curves provide essential information for image

reconstruction of contact and holographic recordings. We found that exposures as small as 2 mJcm^{-2} can produce patterns 20 nm deep. Since the AFM can resolve ~ 2 nm depth differences, about ten vertical or exposure levels are distinguishable. This exposure level is significantly lower than that previously used with other depth measurement techniques.^{1,5} Further improvement should be possible by decreasing the pixel noise, perhaps by modifying the film preparation and developing procedures.

ACKNOWLEDGMENTS

This work was jointly supported by the National Science Foundation, and the U.S. Air Force Office of Scientific Research.

- ¹G. D. Kubiak *et al.*, in *OSA Proceedings on Soft-X-Ray Lithography*, edited by J. Bokor (Optical Society of America, Washington, DC, 1991), Vol. 12, pp. 124–128.
- ²R. J. Hawryluk, H. I. Smith, A. Soares, and A. M. Hawryluk, *J. Appl. Phys.* **46**, 2528 (1975).
- ³J. S. Greeneich, *J. Electrochem. Soc. Solid State Sci. Technol.* **122**, 970 (1975).

- ⁴*X-Ray Microscopy III*, edited by A. G. Michette, G. R. Morrison, and C. J. Buckley (Springer, Berlin, 1992).
- ⁵G. C. Bjorklund, S. E. Harris, and J. F. Young, *Appl. Phys. Lett.* **25**, 451 (1974).
- ⁶M. Howells *et al.*, *Science* **238**, 514 (1987).
- ⁷K. Shinohara, A. Ito, and Y. Kinjo, in *Applications of Laser Plasma Radiation* (SPIE, Bellingham, WA, 1994), Vol. 2015, pp. 10–19.
- ⁸R. A. Cotton *et al.*, in *Applications of Laser Plasma Radiation* (SPIE, Bellingham, WA, 1994), Vol. 2015, pp. 86–96.
- ⁹A. D. Stead, A. M. Page, and T. W. Ford, in *Applications of Laser Plasma Radiation II* (SPIE, Bellingham, WA, 1995), Vol. 2523, pp. 40–50.
- ¹⁰K. Kudo *et al.*, *Jpn. J. Appl. Phys.* **1** **29**, 2572 (1990).
- ¹¹K. Kudo *et al.*, *Jpn. J. Appl. Phys.* **1** **31**, 401 (1992).
- ¹²J. C. White *et al.*, *Appl. Phys. Lett.* **44**, 22 (1984).
- ¹³PMMA 950k Resist in Chlorobenzene, Cat. No. 897631, OCG Microelectronics, 3 Garret Mountain Plaza, West Paterson, NJ 07424.
- ¹⁴Picomotor, New Focus, Inc., 2630 Walsh Avenue, Santa Clara, CA 95051.
- ¹⁵J. A. Samson, *Techniques of Vacuum Ultraviolet Spectroscopy*, 2nd ed. (Pied, Lincoln, NE, 1980).
- ¹⁶AutoProbe Microscope System, Park Scientific Instruments, 1171 Borregas Ave., Sunnyvale, CA 94089.
- ¹⁷G. H. Bernstein, D. A. Hill, and W. Liu, *J. Appl. Phys.* **71**, 4066 (1992).

Design optimization for two lens focused ion beam columns

Li Wang^{a)}

Schlumberger Technologies, ATE, San Jose, California 95110

(Received 13 August 1996; accepted 11 April 1997)

This article systematically analyzes all possible operation modes in a two lens focused ion beam column utilizing analytical formalism and numerical analysis. The non-crossover mode is found to be inherently superior to the crossover mode, and a column in the crossover mode with a single fixed aperture can never realize optimal magnifications for all beam current levels. The non-crossover mode gives the best resolution. The crossover mode, however, offers many attractive features such as simplicity and quick changeover time. Optimization procedures based on operational mode analysis with examples are also given. © 1997 American Vacuum Society. [S0734-211X(97)03204-6]

I. INTRODUCTION

As focused ion beam technology (FIB) is an important tool in the semiconductor industry as well as other scientific and technological fields, the requirements of a FIB system for higher performance become more demanding. To meet these challenges, FIB columns must be optimized under limitations of the liquid metal ion source (LMIS) (energy spread 5–10 eV), the restrictions imposed by pure electrostatic lenses and the complexity of the mechanical construction. The goal of an optimized design is to focus an ion beam to as small a spot as possible for a given beam current. Before a new ion source with a smaller energy spread and an equivalent or higher brightness compared to LMIS such as the gas field ion source,¹ and synthesized lenses² become readily available, we are restricted to using LMIS and lenses with only simple geometry. In practice, two lens FIB columns with three or four element immersion or Einzel lenses are generally employed with significant success. There is still a lot of room to improve the resolution and current densities offered by two lens columns through design optimization. The goal of this study is to provide a systematic, analytical explanation of the advantages and disadvantages of all possible operation modes of a two lens column under justifiable assumptions and to provide optimization procedures based on operation mode analysis.

In the past few years, there has been a number of valuable works on this subject,^{3–6} that have provided much better understanding. Unfortunately, those works contain various limitations or are incomplete. Reference 3 compared the crossover mode shown in Fig. 1(a) with the parallel mode shown in Fig. 1(b), and drew the conclusion that the parallel mode gave smaller beam sizes by showing a numerical example for a fixed total magnification of $M=1$. Although the conclusion is not incorrect, it seems somewhat unfair as the comparison should really be made for a fixed beam current or for an effective aperture size instead of a fixed magnification M . Reference 4 reports on only the crossover mode studied and assumed, without sufficient justification, spherical and chromatic aberration coefficients were not dependent

on M ; thus, the derived expressions may be inaccurate. References 5 and 6 used the thin lens approximation to calculate beam sizes for different modes and took Coulomb interactions into account. However, the authors failed to provide systematic analytical reasoning, and also did not point out the dependence of effective aperture size on total magnification.

Based on the author's own calculations using a commercial software package,⁷ Coulomb interactions seem to be insignificant for typical two lens columns when the beam current is small (<100 pA) and when interactions between the LMIS and the first beam acceptance aperture are included in the virtual source size (about 50 nm). Therefore, this study does not include Coulomb interactions which are considered by some studies^{5,6} as an important factor limiting the resolution of FIB systems. Further discussions on this subject are given in Section IV.

This article is organized as follows. In Section II the author uses the thin lens approximation to deduce the conclusion that the non-crossover mode always gives a smaller beam size than the crossover mode for a given beam current if both are in their respective optimal magnifications. In Section III, some practical implementations of the optimization on all the modes in Fig. 1 are discussed by using some realistic numerical examples. In Section IV, the effects of thick lenses and Coulomb interaction are discussed. Final conclusions are given in Section V.

II. WHICH MODE IS THE BEST?

The performance of a charged particle optical system is frequently characterized in terms of the spot size of the focused beam. Here the beam size d is evaluated by the commonly used quadrature sum of the contributions due to the Gaussian image of the source size d_g , spherical aberration d_s and chromatic aberration d_c as follows:

$$d = \sqrt{d_g^2 + d_s^2 + d_c^2}, \quad (1)$$

where

$$d_g = M d_o, \quad d_s = \frac{1}{2} M \alpha^3 C_{so}, \quad d_c = M \alpha C_{co} \Delta E / E.$$

^{a)}Electronic mail: liwang@san-jose.ate.slb.com

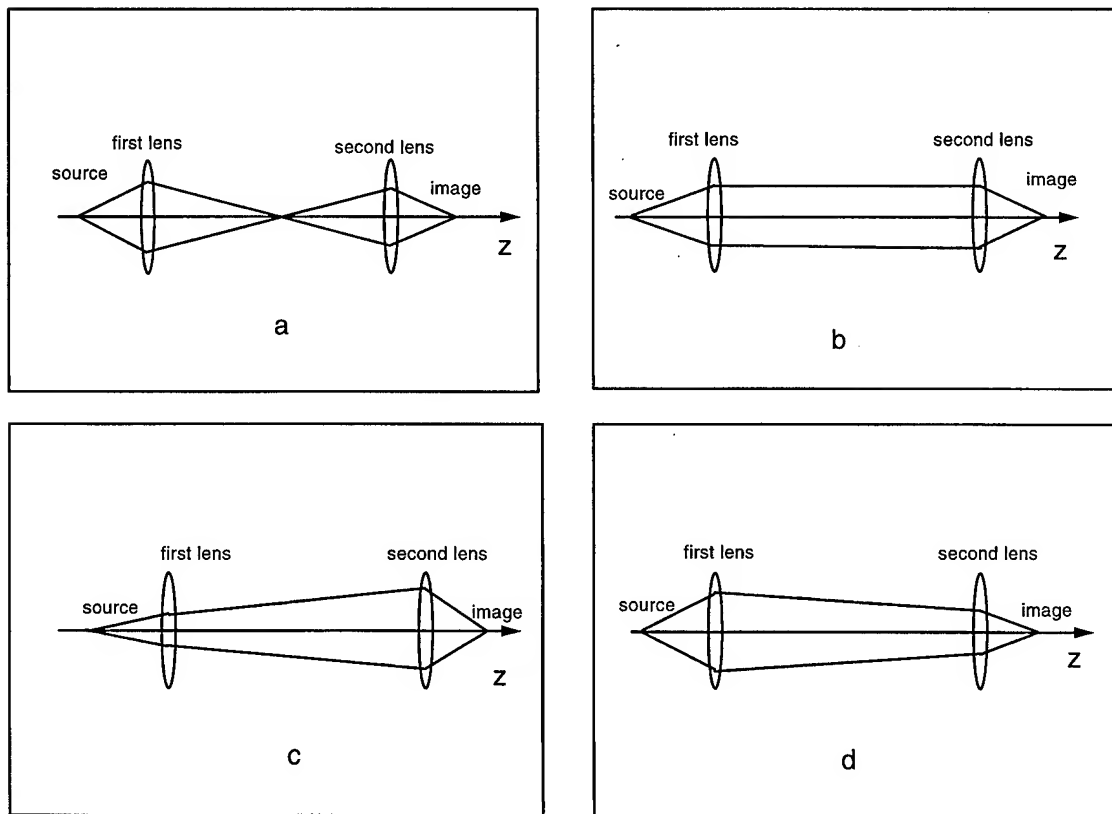


FIG. 1. Operational modes of a two lens FIB column: (a) crossover mode, (b) parallel mode, (c) diverging mode, and (d) converging mode where (b), (c), and (d) are the non-crossover modes.

M is the total magnification of the column; α is the half-angle subtended by the effective beam-limiting aperture; d_v is the virtual source size; C_{so} is the total spherical aberration coefficient that refers to the object side of the column; C_{co} is the total chromatic aberration coefficient that refers to the object side of the column; ΔE is the energy spread of the ion source while E is the beam energy. The deflection and diffraction aberrations are not considered here since most of the deflection aberrations can be eliminated through dynamic corrections⁸ and the scanning area is very small when a FIB

system is operated at a very high magnification. In addition, the de Broglie wavelength of the ions is extremely small in the current FIB energy range.

For a two lens column, the total spherical and chromatic aberration coefficients can be expressed in terms of the aberration coefficients of the two lenses:⁹

$$C_{so} = C_{s1} + C_{s2}g^3/M^4, \quad (2)$$

$$C_{co} = C_{c1} + C_{c2}g^3/M^2, \quad (3)$$

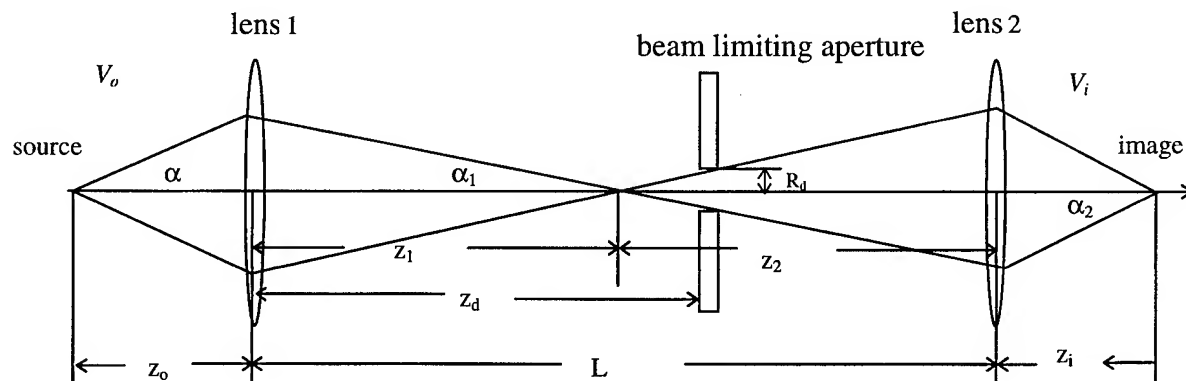


FIG. 2. Definitions of the optical parameters in a two lens FIB column.

where C_{s1} is the spherical aberration coefficient of the first lens (condenser) that refers to the object side; C_{s2} is the spherical aberration coefficient of the second lens (objective) that refers to the image side; C_{c1} is the chromatic aberration coefficient of the first lens (condenser) that refers to the source side; C_{c2} is the chromatic aberration coefficient of the second lens (objective) that refers to the target side; $g = \sqrt{V_o/V_i}$ where V_o and V_i are the acceleration voltages in the source side and sample side respectively.

Normally, the two lenses used in a FIB column can be viewed as thin lenses without introducing a lot of errors.¹⁰ Under the thin lens approximation, there is no difference between the real and asymptotic optical parameters, and C_{s1} , C_{s2} , C_{c1} , and C_{c2} can be represented by m_1 and m_2 (the magnifications of the first and second lenses respectively) as follows:⁹

$$C_{s1} = C_{s1\infty}(1 - 1/m_1)^4, \quad C_{s2} = C_{s10}(1 - m_2)^4,$$

$$C_{c1} = C_{c1\infty}(1 - 1/m_1)^2, \quad C_{c2} = C_{c20}(1 - m_2)^2,$$

where $C_{s1\infty}$ and C_{s20} are the spherical aberration coefficients on the object side of the first lens with an infinite magnification and on the image side of the second lens with a zero magnification respectively, while $C_{c1\infty}$ and C_{c20} are the chromatic aberration coefficients on the object side of the first lens with an infinite magnification and on the image side of the second lens with a zero magnification respectively. Since $m_1 = M/m_2$, and substituting it into Eqs. (1)–(3), we have

$$d = \sqrt{(Md_v)^2 + \left\{ 0.5M\alpha^3 \left[C_{s1\infty} \left(1 - \frac{m_2}{M} \right)^4 + C_{s20} \frac{(1 - m_2)^4}{M^4} \right] \right\}^2 + \left\{ M\alpha \frac{\Delta E}{E_0} \left[C_{c1\infty} \left(1 - \frac{m_2}{M} \right)^2 + C_{c20} \frac{(1 - m_2)^2}{M^2} \right] \right\}^2}. \quad (4)$$

When we optimize the column, we must first fix some parameters, then finish the optimization process by varying other parameters. Otherwise, we would have too many variables to worry about at the same time. For a two lens column shown in Fig. 2, the usual situation would be that the z_o , z_i , V_o , V_i , and L should be easily decided earlier from practical considerations while other things such as the aberration coefficients and the aperture position are being decided. The optimization process to be discussed here is how to find the best magnification, aperture size and position, etc. not for individual lens optimization.

In the following discussion, we use the sign rule for the optical parameters: we define z_o , z_i , V_o , V_i , and L as positive; z_1 is positive if the beam crosses the axis after the first lens [crossover mode in Fig. 1(a) and converging mode in Fig. 1(d)] while z_1 is negative if the beam has a virtual crossover on the optical axis [diverging mode Fig. 1(c)]; if the slope of the angle subtended by the ion trajectories and the positive z direction is positive, then this angle is positive, while the angle is negative when its slope is negative.

Under this sign rule, it is easy to derive the following relations $m_1 = -(z_1/z_o)g$, $m_2 = -z_i/z_2$ and

$$M = m_1^* m_2 = \frac{z_i}{z_o} \frac{z_1}{z_2} g, \quad (5)$$

also

$$\frac{\alpha_1}{\alpha} = \frac{g}{m_1} = \frac{z_o}{z_1}. \quad (6)$$

We have assumed that the second lens is an Einzel lens for simplicity considering the fact that the objective lens is a unipotential lens in the majority of practical FIB columns. If

the objective lens is not an Einzel lens such as the retarding type in the low energy deposition system, all the relations in this article are still valid provided that z_i is replaced by $z_i\sqrt{V_1/V_i}$, where V_1 is the acceleration voltage in the image position of the first lens.

Since $z_2 = L - z_1$, we have $z_1 = LMz_o/(gz_i + Mz_o)$. Then

$$m_2 = -\frac{z_i}{L - z_1} = -\frac{gz_i + Mz_o}{Lg}. \quad (7)$$

Therefore, once z_i , z_o and L are known, m_2 and M are interdependent.

Although the above expressions are derived based on Fig. 2 which shows a crossover mode, all the formulas are valid for non-crossover modes such as in Figs. 1(b)–1(d). The difference is that we have $z_1 < -z_o$ for the diverging mode of Fig. 1(c), $z_1 > L$ for the converging mode of Fig. 1(d), and $z_1 = \infty$ for the parallel mode. The case $-z_o < z_1 < 0$ is not real since divergent lenses do not exist in charged particle optics. For the diverging mode, since $z_1 < -z_o$ and $z_2 = L - z_1 > 0$, then $m_1 > 0$, $m_2 < 0$ and $M < 0$. For the converging mode, since $z_1 > L > 0$ and $z_2 = L - z_1 < 0$, then $m_1 < 0$, $m_2 > 0$ and $M < 0$. For the parallel mode, $m_1 = \infty$, $m_2 < 0$, and $M = -z_i g/z_o < 0$. In other words, $M < 0$ is always true for the non-crossmodes. For the crossover mode, $m_1 < 0$, $m_2 < 0$ thus, we have $M > 0$. Also, we can see $|M_{\text{diverging}}| < |M_{\text{parallel}}| < |M_{\text{converging}}|$. Therefore, we know which mode it is just by determining the sign and the magnitude of M .

Now we are in a position to discuss which mode gives smallest d for a given α (equivalent to a given beam current I_b) and for a given set of V_o , V_i , z_o , z_i , and L . Substituting Eq. (7) into the expression of d_c we have

$$d_c = \left(\alpha \frac{\Delta E}{E_0} \right) \left[C_{c1\infty} M \left(1 + \frac{gz_i + Mz_i}{LgM} \right)^2 + C_{c20} \left(1 + \frac{gz_i + Mz_o}{Lg} \right)^2 / M \right] \\ = \left(\alpha \frac{\Delta E}{E_0} \right) \frac{a_2 M^2 + a_1 M + a_0}{L^2 g^2 M}, \quad (8)$$

where

$$a_0 = C_{c20} L^2 g^5 + 2 C_{c20} L g^5 z_i + C_{c1\infty} g^2 z_i^2 + C_{c20} g^5 z_i^2,$$

$$a_1 = 2 C_{c1\infty} L g^2 z_i + 2 C_{c20} L g^4 z_o + 2 C_{c1\infty} g z_i z_o \\ + 2 C_{c20} g^4 z_i z_o,$$

$$a_2 = C_{c1\infty} L^2 g^2 + 2 C_{c1\infty} L g z_o + C_{c1\infty} z_o^2 + C_{c20} g^3 z_o^2,$$

and a_0 , a_1 , a_2 are all positive numbers. In this article $C_{s1\infty}$, C_{s20} , $C_{c1\infty}$, and C_{c20} are assumed to be independent of M . To find the minimum d_c , we need to solve $d[d_c]/dM=0$, i.e.,

$$\frac{d[d_c]}{dM} = \left(\alpha \frac{\Delta E}{E_0} \right) \frac{a_2 M^2 - a_0}{L^2 g^2 M^2} = 0.$$

The solutions are $M=M_c$ and $M=-M_c$ where $M_c = \sqrt{a_0/a_2}$. Since $d_c(M_c) = (\alpha \Delta E/E_0)(a_2 M_c^2 + a_1 M_c + a_0)/L^2 g^2 M_c > d_c(-M_c) = (\alpha \Delta E/E_0)(a_1 M_c - a_2 M_c^2 - a_0)/L^2 g^2 M_c$ and $M < 0$ means the non-crossover modes, while $M > 0$ means the crossover mode, we conclude that, between the two local minimums, the non-crossover minimum is the global minimum.

Now let us proceed to discuss d_s . Substituting Eq. (7) into the expression of d_s we have

$$d_s = 0.5 \alpha^3 \frac{b_4 M^4 + b_3 M^3 + b_2 M^2 + b_1 M + b_0}{L^4 g^4 M^3}, \quad (9)$$

where

$$b_0 = C_{s20} L^4 g^7 + 4 C_{s20} L^3 g^7 z_i + 6 C_{s20} L^2 g^7 z_i^2$$

$$+ 4 C_{s20} L g^7 z_i^3 + C_{s1\infty} g^4 z_i^4 + C_{s20} g^7 z_i^4,$$

$$b_1 = 4 C_{s1\infty} L g^4 z_i^3 + 4 C_{s20} L^3 g^6 z_o + 12 C_{s20} L^2 g^6 z_i z_o$$

$$+ 12 C_{s20} L g^6 z_i^2 z_o + 4 C_{s1\infty} g^3 z_i^3 z_o + 4 C_{s20} g^6 z_i^3 z_o,$$

$$b_2 = 6 C_{s1\infty} L^2 g^4 z_i^2 + 12 C_{s1\infty} L g^3 z_i^2 z_o + 6 C_{s20} L^2 g^5 z_o^2$$

$$+ 12 C_{s20} L g^5 z_i z_o^2 + 6 C_{s1\infty} g^2 z_i^2 z_o^2 + 6 C_{s20} g^5 z_i^2 z_o^2,$$

$$b_3 = 4 C_{s1\infty} L^3 g^4 z_i + 12 C_{s1\infty} L^2 g^3 z_i z_o + 12 C_{s1\infty} L g^2 z_i z_o^2$$

$$+ 4 C_{s20} L g^4 z_o^3 + 4 C_{s1\infty} g z_i z_o^3 + 4 C_{s20} g^4 z_i z_o^3,$$

$$b_4 = C_{s1\infty} L^4 g^4 + 4 C_{s1\infty} L^3 g^3 z_o + 6 C_{s1\infty} L^2 g^2 z_o^2$$

$$+ 4 C_{s1\infty} L g z_o^3 + C_{s1\infty} z_o^4 + C_{s20} g^3 z_o^4.$$

Obviously, b_0 , b_1 , b_2 , b_3 and b_4 are all positive numbers. Let us suppose $M=M_s > 0$ gives the smallest d_s for all $M > 0$, then we can conclude that $M=-M_s$ gives even a smaller value of d_s since

$$d_s(M_s) = 0.5 \alpha^3 \frac{b_4 M_s^4 + b_3 M_s^3 + b_2 M_s^2 + b_1 M_s + b_0}{L^4 g^4 M_s^3} \\ > 0.5 \alpha^3 \frac{b_3 M_s^3 + b_1 M_s + b_0 - b_2 M_s^2 - b_4 M_s^4}{L^4 g^4 M_s^3} \\ = d_s(-M_s).$$

Thus, $M=M_s$ must not be the global minimum. In addition, unlike the d_c situation, $M=-M_s$ is not necessarily the minimum position for the non-crossover modes. In other words, the global minimum value of d_s could be even less than $d_s(-M_s)$.

As a matter of fact, for an arbitrary $M=M_0 > 0$, from Eq. (8) and Eq. (9), we have the relation $d^2(M_0) = M_0^2 d_v^2 + d_c^2 + d_s^2 > d^2(-M_0)$ being always true by using similar reasoning. Therefore, we can safely conclude that the non-crossover mode can always give better performance than the crossover mode since if $M=M_{\min} > 0$ gives the smallest d for all $M > 0$, there is always a non-crossover mode with a magnification $-M_{\min}$ such that $d(-M_{\min}) < d(M_{\min})$ and, furthermore, $d(M < 0)$ could be smaller than $d(-M_{\min})$.

Since $|M_{\text{diverging}}| < |M_{\text{parallel}}| < |M_{\text{converging}}|$, and in general, for smaller α , the optimal M is smaller too, while for larger α , larger M is preferred for the best performance. Thus, usually for a very small beam current the diverging mode offers the best results, and for very large current the converging mode is better. Of course, for a specific column, the best M depends on the specific setup.

III. OPTIMIZATION PROCEDURE

The beam current is related to the effective aperture as follows. $I_b = \pi (dI/d\Omega) \alpha^2$, where $dI/d\Omega$ is the angular current density of the LMIS. What we want in an optimization procedure is to obtain the optimized, realizable I_b - d curve. This curve will serve as a guide to a practical design of a FIB column. Obviously, we should first optimize the individual lenses by taking all the mechanical and economical limitations into account. Then we should find the best magnification for a given I_b . Finally, we should study the requirements of the optimized M on the practical setup and come up with the best size and position of the beam limiting aperture, after some compromises if necessary. But optimization of lenses is not the intent of this study as we indicated earlier. Instead, we are going to concentrate on the last two steps.

To find the best M for a given set of aberration coefficients, acceleration voltages, minimum working distances and total length of the column is equivalent to solving

$$\frac{d[d^2(M)]}{dM} = 0. \quad (10)$$

Substituting Eq. (8) and Eq. (9) into Eq. (10), we will end up with a polynomial equation with a form like

$$c_0 + c_1 M + c_2 M^2 + c_3 M^3 + c_4 M^4 + c_5 M^5 + c_6 M^6 \\ + c_7 M^7 + c_8 M^8 = 0, \quad (11)$$

TABLE I. The assumed parameters of the LMIS and the two lenses used in the examples.

d_v	V_o	V_i	E	ΔE	$C_{s1\infty}$	$C_{c1\infty}$	C_{s20}	C_{c20}	$dI/d\Omega$	L	z_o	z_i
50 nm	7.2 keV	30 keV	30 keV	5 eV	400 mm	50 mm	2000 mm	150 mm	17 $\mu A/str$	300 mm	20 mm	20 mm

where c_n ($n=0, \dots, 8$) are expressed in terms of d_0 , α , a_0 , a_1 , a_2 , and b_0 , b_1 , b_2 , b_3 , b_4 . c_n ($n=0, 1, \dots, 8$) are too lengthy (but elementary) to list them here. In general, we will get as many as eight solutions mathematically. But most of them are not real. According to the author's own experience, normally only two roots are real and one is positive (corresponds to the crossover mode) while the other is negative (corresponds to the non-crossover mode).

Let us consider a numerical example here. The values of the defining parameters for the two lens FIB column being discussed are listed in Table I. The aberration coefficient values are typical compared to many commercially and non-commercially available columns using deceleration mode lenses. Of course, the values could be much better if lenses with acceleration modes are used.¹¹

By solving the Eq. (11) for this example, we find that, just as we expected, two real solutions with opposite signs exist for every value of α . The solutions are plotted in Fig. 3. Using those values of magnification, d , d_c , d_s , and d_g for the noncrossover and crossover modes are plotted in Figs. 4–6. Clearly, just as we expected, Fig. 4 shows that the non-crossover mode is always better than the crossover mode if both modes are operated in their respective optimal magnifications for all beam current levels. Figure 5 and Fig. 6 show the variations of the dominance by d_c , d_s , and d_g as the beam current changes are similar in the two modes. Also, as we mentioned in Section II, for the lower current range, which is $I_b = 2$ pA–250 pA for this case, the diverging mode gives optimal performance, while the converging mode offers the smallest beam sizes for higher current ranges ($I_b > 250$ pA).

So far we have discussed how to find the optimal magnification for each I_b . There is still an important question to

answer: Can these optimal magnification values be realized in a practical column? The answer is very different for columns using crossover modes and those using non-crossover modes.

A. Aperture selection for crossover mode columns

We have shown that the crossover mode is always naturally inferior to the noncrossover mode if both have their optimum conditions of smallest beam sizes. Crossover mode columns do have some advantages over non-crossover mode columns, however, such as much simpler mechanical structures, much quicker current changeover time, better reliability and serviceability and a lower cost. When the targeted application does not require superb resolutions, crossover mode columns may be the better solution. Therefore it is of great practical importance to study how to optimize a crossover mode column.

In practice, the crossover mode is realized by a single aperture whose size and position are fixed as shown in Fig. 2. The beam current I_b , or the effective aperture α , is adjusted by changing the first lens strength. This type of beam limiting aperture is usually called an electrically variable aperture (EVA). Let the radius of the EVA be R_d and the distance between the first lens and the EVA position be z_d . Figure 2 shows us $|\alpha_1| = \arctan(R_d/|z_d - z_1|)$. From Eqs. (6) and (7), we have the following relation:

$$\alpha = \frac{LM}{gz_i + Mz_o} \arctan\left(\frac{R_d(gz_i + Mz_o)}{z_d(gz_i + Mz_o)LMz_o}\right). \quad (12)$$

Therefore, once R_d and z_d are fixed, α and M are not independent of each other, i.e., we can never operate an EVA

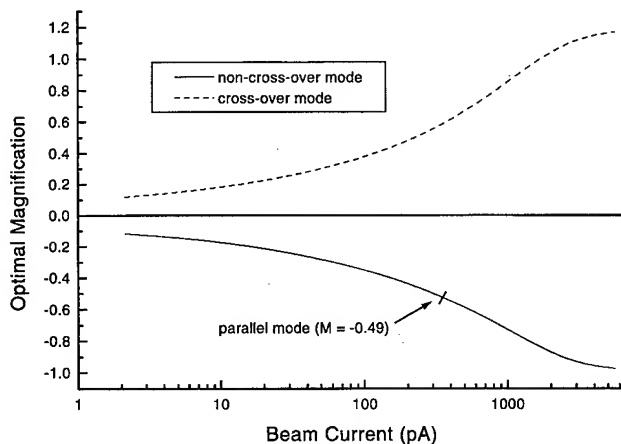


FIG. 3. The optimal magnifications for the crossover mode and the non-crossover mode for the column defined in Table I.

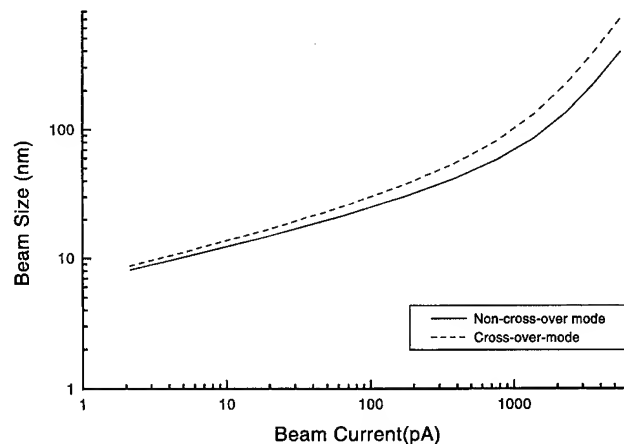


FIG. 4. A comparison when both the crossover mode and the non-crossover mode are in their optimal magnifications.

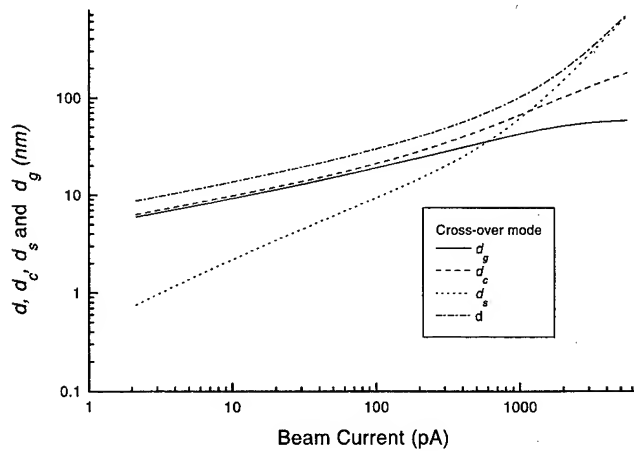


FIG. 5. The best possible beam size d , Gaussian image d_g , chromatic aberration d_c , and spherical aberration d_s as functions of beam current I_b for the crossover mode.

column in optimal magnifications for all current levels. What we can do is optimize M for the smallest I_b for the highest imaging resolutions, optimize M for larger I_b for better micromachining results and speeds, or compromise somewhere in between depending on the specific situation.

To illustrate this point further, the I_b versus the required R_d for the best magnifications of $z_d=100$ mm, 150 mm, 200 mm, and 250 mm in the column described by Table I are plotted in Fig. 7. Figure 7 shows that $z_d=250$ mm is the best choice of the four EVA positions since this curve is much smoother and flatter than the other three, which means the difference between the best possible I_b - d curve and the I_b - d curve with this particular choice of R_d would be smaller. Figure 8 shows the corresponding I_b - d curves when R_d is chosen for optimized beam sizes of $I_b=2$ pA, 64 pA, and 1389 pA when $z_d=250$ mm. The optimized R_d 's are 13 μm , 24 μm and 27 μm respectively. The curve for best 2 pA performance degrades the beam size significantly for most I_b levels except the very small I_b range; the curve optimized for 1389 pA improves for larger I_b by compromising the

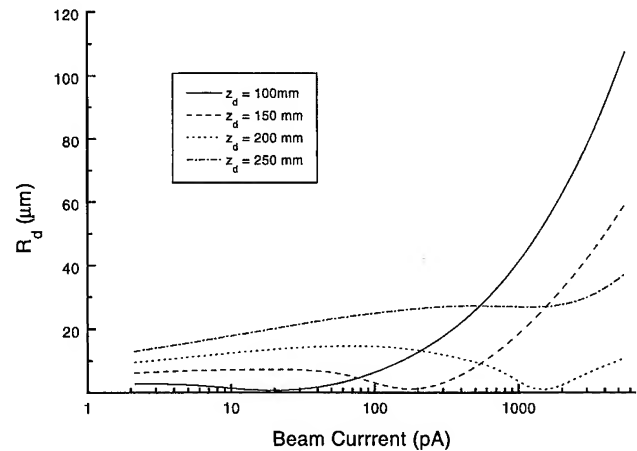


FIG. 7. The radius of the beam limiting aperture R_d as a function I_b for the aperture position $z_d=100$ mm, 150 mm, 200 mm, and 250 mm in the cross-over mode.

beam sizes in the smallest I_b range. The curve optimized for 64 pA falls between the two, offering a good balance for all I_b ranges. Note that the best possible I_b - d curve is also plotted in Fig. 8 for comparison. We can see that the difference among the optimized R_d , targeting different I_b 's, does not differentiate the overall beam sizes very much, although it is still noticeable because of the flatness of the I_b - R_s curve for $z_s=250$ mm as shown in Fig. 7. Another factor to be considered here is the practical machining and reliability of the size of EVA required. The conclusion is that a 24 μm aperture is preferable because machining errors as well as alignment errors would be smaller for a 24 μm aperture than those for a 10 μm aperture. In addition, the aperture lifetime would be longer for the larger aperture.

B. Aperture selection for non-crossover mode columns

In practice, the current change of a non-crossover mode column is realized by a mechanical variable aperture (MVA) which is a plate with a fixed number of various size holes. In

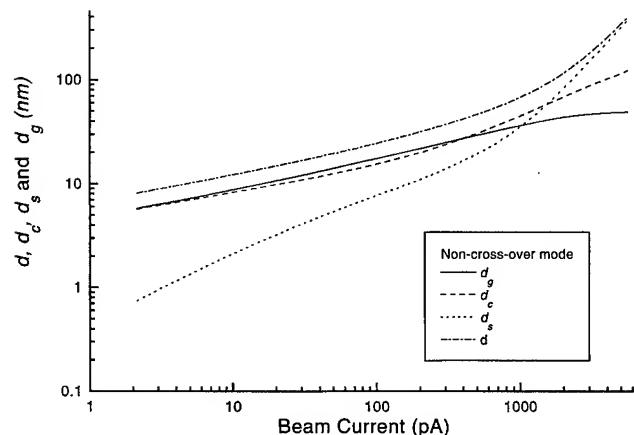


FIG. 6. The best possible beam size d , Gaussian image d_g , chromatic aberration d_c , and spherical aberration d_s as functions of beam current I_b for the non-crossover mode.

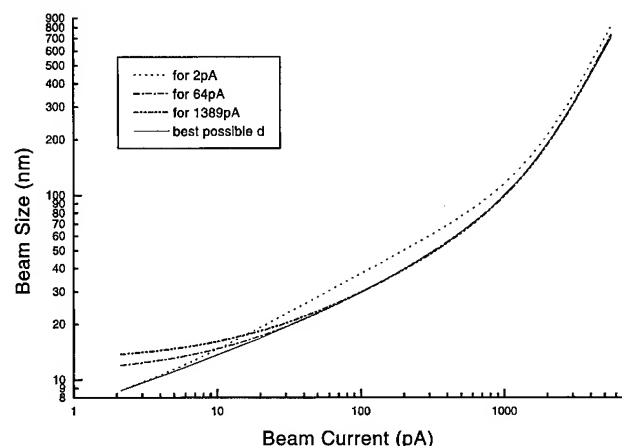


FIG. 8. The beam size d optimized for $I_b=2$ pA, 64 pA, and 1389 pA and the best possible beam size as a function of I_b .

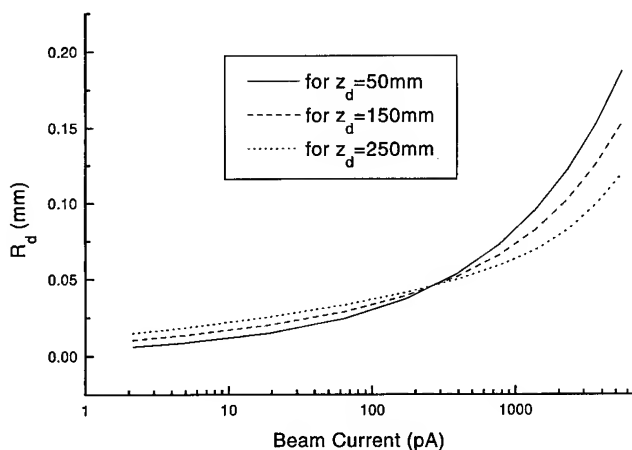


FIG. 9. The radius of the beam limiting aperture R_d as a function I_b for the aperture position $z_d = 50$ mm, 150 mm, and 250 mm in the non-crossover mode.

other words, z_d is fixed while R_d can be changed noncontinuously. The answer to the question whether or not the best M for all I_b can be materialized is generally yes in this case since, according Eq. (12), by varying R_d any value of M for a given α can be provided. We now must calculate the corresponding MVA sizes for the best M and the best location in which to install it.

By substituting the best magnifications we calculated earlier for various I_b 's to Eq. (12) we can calculate the required R_d for any chosen z_d . For the column we described in Table I, the calculated R_d 's are plotted in Fig. 9 for $z_d = 50$ mm, 150 mm and 250 mm. The choice of z_d should be based on the convenience of the total column construction and its reliability. If $z_s = 50$ mm, then the corresponding optimized $R_d = 6 \mu\text{m}$ for 2 pA is a little bit too small for machining and has a possible shorter lifetime due to ion beam etching. Therefore, we should go for a larger R_d , e.g., $R_d = 10 \mu\text{m}$ and $14 \mu\text{m}$ for $z_d = 150$ mm and 250 mm, respectively, for the 2 pA beam unless there are other constraints involved.

IV. DISCUSSIONS

As mentioned previously, Coulomb interactions have not been taken into account since there is no widely accepted theory with good quantitative accuracy yet available, although there has been some progress in this direction.⁶ It is difficult to judge the accuracy of the results given in these studies, since many assumptions and simplifications were used because of the extremely complicated nature of Coulomb interactions. This author has calculated Coulomb interaction effects in some typical two lens FIB column configurations using the commercial software package, Boersch. The results obtained show that beam size increases due to Coulomb interaction for small beam currents are in the range of 0.5 – 2 nm, which is significantly less than the quadrature sum of other contributions. Coulomb interactions do become a significant factor, however, when the beam currents are

over several hundred pA. We believe that the analysis given in this study is appropriate for small and medium beam currents (up to several hundred pA) which seem to be the most useful range for FIB applications. On the other hand, for larger currents of a few nA, beam sizes predicted by our study may be too optimistic because Coulomb interactions could play a major role. Since most Coulomb interaction theories indicate that crossovers enlarge the effects of Coulomb interactions, we can believe that the non-crossover mode will give better resolution performance than the crossover mode in all cases.

Another basic approximation used in this article is the thin lens approximation. Even though we have a number of reasons to believe that this approach can be justified in many cases, there is a possibility, however, that a thick lens model must be adopted in some instances. The effect of the thick lenses will be a topic of future studies.

V. CONCLUSIONS

The non-crossover mode can always outperform the crossover mode if both are operating at optimal conditions, and the diverging mode suits smaller beam current levels whereas the converging mode is better for larger currents. This conclusion was deduced based on analytical reasoning rather than on specific numerical examples. The best possible magnifications for the whole beam current range cannot be realized in a crossover mode column with an EVA. On the other hand, the non-crossover mode with a MVA can always realize its smallest possible predicted theoretically beam sizes. An optimization procedure with adequate numerical discussions for all operation modes was also shown. Since Coulomb interaction effects are not included in this study, the analysis shown in this article may not be suitable for all column configurations and beam current ranges.

ACKNOWLEDGMENTS

The author would like to thank Dr. T. Lundquist and D. Keckley of Schlumberger for their encouragement and support, and Dr. J. Li of FEI Company and X. Jiang of Delft University of Technology for their helpful discussions. Finally, the author would like to express his gratitude to Dr. J. Orloff of the University of Maryland for his constant support in the past few years.

¹K. Jouston, K. Bohringer, and S. Kalbitzer, *Appl. Phys. B* **46**, 313 (1988).

²M. Szilagyi, *J. Vac. Sci. Technol. B* **1**, 1137 (1983).

³J. Orloff, *J. Vac. Sci. Technol. B* **5**, 175 (1987).

⁴Y. L. Wang and Z. Shao, *Adv. Electron. Electron Phys.* **81**, 177 (1991).

⁵X. R. Jiang and P. Kruit, *Microelectron. Eng.* **30**, 249 (1996).

⁶P. Kruit and X. R. Jiang, *J. Vac. Sci. Technol. B* **4**, 1635 (1996).

⁷"Boersch," a software package provided by Munro's Electron Beam Software Ltd.

⁸E. Munro, *Charged Particle Optics B*, edited by A. Septier (Academic, New York, 1980), p. 74.

⁹M. Szilagyi, *Electron and Ion Optics* (Plenum, New York, 1988).

¹⁰T. Ishitani and Y. Kawanami, *J. Vac. Sci. Technol. B* **13**, 371 (1995).

¹¹K. Kurihara, *J. Vac. Sci. Technol. B* **3**, 41 (1985).

Field emission from amorphous diamond coated Mo tip emitters by pulsed laser deposition

M. Q. Ding,^{a)} A. F. Myers, W. B. Choi, R. D. Vispute, S. M. Camphausen, J. Narayan, J. J. Cuomo, and J. J. Hren

Department of Materials Science and Engineering, North Carolina State University, Raleigh, North Carolina 27695-7907

J. Bruley

Department of Materials Science and Engineering, Lehigh University, Bethlehem, Pennsylvania 18105

(Received 8 November 1996; accepted 18 April 1997)

Previous studies have shown that carbon films deposited on needle-shaped Si emitters by filtered cathodic arc are amorphous with a high sp^2 content. These results can be ascribed to the poor thermal conductivity inherent to this geometry. Our present studies overcome this difficulty by depositing amorphous diamond films on Mo tip emitters by pulsed laser deposition. Monitoring films were grown on sapphire substrates and appeared transparent with a resistivity greater than $1 \times 10^6 \Omega \text{ cm}$, showing a typical amorphous diamond nature. Electron energy loss spectroscopy showed that the sp^3 content of the film was 50% at the apex of the tip and 30% at the shank, which was lower than on planar substrates. High resolution transmission electron microscopy images revealed that the film at the apex was much denser than that at the shank, but both film showed a nano-columnar microstructure. Selected area electron diffraction confirmed that the films were amorphous in nature. Field emission from coated Mo tip emitters showed a considerable improvement in both turn-on voltage and emission current, compared with the same emitter before deposition. It has been demonstrated for the first time, that a-D films can be deposited on Mo tip emitters by pulsed laser deposition to achieve a better emission performance and the observation of this in the as-deposited structure. © 1997 American Vacuum Society. [S0734-211X(97)06204-5]

I. INTRODUCTION

Field emitters coated with diamond films have shown a great improvement in performances, such as low onset voltage, high current density, and better stability.¹⁻⁸ They are ascribed to the fact that diamond films have low or even negative electron affinity and chemical inertness.^{9,10} On the other hand, diamondlike carbon (DLC, hydrogenated) and amorphous diamond (a-D or ta-C, nonhydrogenated) films have also shown attractions to field emission, since these films retain similar properties to diamond, such as extreme hardness, high density, high resistivity, chemical inertness, etc.,¹¹⁻¹⁷ due to the high percentage of sp^3 bonding in DLC and a-D films. They also contain a fraction of sp^2 bonding embedded in the sp^3 network, which form nonlocalized π states and optical energy gaps ranging from 0 to 2.7 eV. However, the properties of DLC are unstable when heated to 300–500 °C due to the release of H_2 leading to conversion of sp^3 to sp^2 , whereas a-D films have been reported stable up to 800 °C.¹⁸ Ion beam sputtering, filtered cathodic arc and laser ablation are the most commonly used techniques for depositing amorphous diamond films.^{11,12,16-19}

In terms of applications of electron devices, such as flat panel displays, a-D films have some advantages over diamond films. First, a-D films are microscopically smooth while diamond films usually are not. Second, a-D films are deposited at room temperature by a rapidly quenching pro-

cess; however, growth of diamond films requires a high substrate temperature (above 650 °C). Third, the preparation of a-D films is simpler and faster than that of diamond films.

One of the applications of DLC and a-D films is as a coating for field emitters. Xie *et al.*²⁰ deposited a-D film on flat Si wafers, by a Q-switched Nd-YAG laser. They found an onset voltage of field emission to be less than 20 V/ μm and a current density of up to 100 mA/ cm^2 . Amaratunga *et al.*²¹ reported field emission also from Si wafers, coated with N-containing hydrogenated amorphous carbon films by capacitor coupled chemical vapor deposition (CVD); onset emission fields as low as 4 V/ μm were observed. Recently, Chuang *et al.*²² reported enhancement of field emission from Mo tip arrays with DLC films; gated Mo tips were coated in a pressure of 10^{-2} mbar (3 mol % H_2 in N_2) by laser ablation. The maximum anode current, however, decreased during operation, possibly due to instability of DLC films. In addition, they were unable to characterize the film on the tip arrays.

Previous studies in this lab²³ have shown that field emission from Si single tips coated with carbon films at room temperature by filtered cathodic arc was reduced as compared to that of the uncoated tip, most likely because the coating was amorphous carbon as characterized by electron energy loss spectroscopy (EELS). In fact, filtered cathodic arc deposition has produced a-D films with a sp^3 content of 93% on planar substrates such as Si, Mo, etc.¹⁷ However, for a sharpened tip with a poor thermal conduction, the heat generated under intensive carbon ion bombardment could

^{a)}On leave from Beijing Vacuum Electronics Research Institute, P. O. Box 749, Beijing 100016, China; Electronic mail: ding@mte.ncsu.edu

raise the temperature of the tip. Therefore, the deposition condition would be no longer in favor of a rapidly quenching process to form sp^3 bonding, resulting in an amorphous carbon film.

So far there have been few reports on field emission from a-D coated Mo tips, though several investigators have reported field emission from DLC coated Si and Mo tip arrays.^{22,24} In the present article, we report field emission studies of single Mo tip emitters coated with a-D films by pulsed laser deposition. With direct EELS and high resolution transmission electron microscopy (HRTEM) studies, we were able to examine true bonding and microstructure of the tip. Scanning electron microscopy (SEM) was employed to study the microscopic morphology of the a-D film, while selected area electron diffraction (SAED) was used to determine its amorphous nature. Field emission from the Mo emitters before and after deposition of a-D films was measured using a field ion microscopy/field emission microscopy (FIM/FEM) system.

II. EXPERIMENTAL PROCEDURES

A. Mo tip preparations

Single needle-shaped emitter tips were made from 99.95% pure 0.125-mm-diam Mo wires. A short length of the wire was mounted in a copper tube and then etched electrolytically in a concentrated potassium hydroxide solution at a dc voltage of 10 V applied between the Mo wire and a Pt electrode. The radii of the emitter tips were usually around 30 nm.

B. Laser ablation of a-D films

Deposition of a-D films was performed in a stainless steel vacuum chamber at a base pressure of 7.5×10^{-6} Pa. A laser beam with a wavelength of 248 nm and a pulse duration of 20 ns full width half-maximum (FWHM) was delivered from a KrF excimer laser. It was focused and hit the target at an angle of 45° to the surface normal; a power density of 25 J/cm² and a repetition rate of 10 Hz were used in the experiment. The high purity graphite target was kept rotating at 10 rpm during the ablation. The Mo tips were mounted about 5 cm away from and pointed towards the target. The substrates were kept at room temperature through the experiment. The deposition process lasted for 10 min, resulting in a film thickness of about 1000 Å.

C. Tip characterizations

The morphology of the a-D films was studied by scanning electron microscopy (JEOL 6400 field emission SEM). HR-TEM and SAED were performed using a Topcon 002B TEM. The film's bonding structure was determined by EELS (VG H8 501 STEM).

D. Field emission measurements

The current-voltage (I-V) characteristics of the Mo tip emitters were studied in an ultrahigh vacuum (UHV) cham-

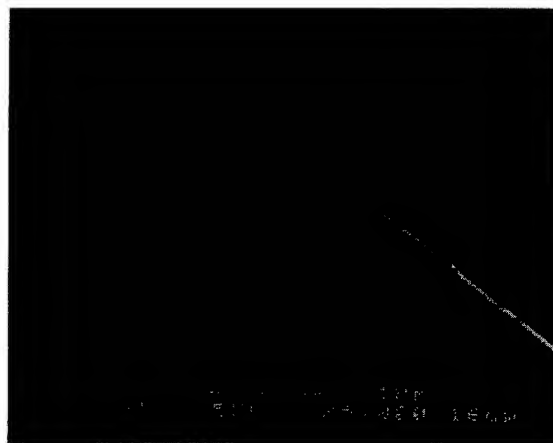


FIG. 1. SEM image of a Mo tip emitter coated with the a-D film, which is continuous and microscopically smooth.

ber using an FIM/FEM system. The Mo tip emitter was placed 10 mm (precision of ± 0.03 mm) away from the anode (Faraday cup). All the measurements were made at room temperature under a vacuum of below 1.33×10^{-7} Pa. Field emission data were collected before and after coating on the same Mo tip, so that field emission performance can be compared accurately. The emission current was measured by a Keithley 617 electrometer with a current sensitivity of 10^{-15} A. The data was acquired by a computer through an IEEE-488 interface with each data point sampled 100 times at a time interval of 0.25 s.

III. RESULTS

A. Properties of a-D films on planar sapphire substrates

To assess laser deposition conditions, monitoring films were deposited on planar sapphire samples. These films were transparent and extremely hard, showing typical characteristics of amorphous diamond. Resistivity of the films was measured by a four-probe method to be greater than 1×10^6 Ω cm. However, these properties only serve as an indication that the deposition conditions were appropriate for the planar samples.

B. Microscopy observations

Figure 1 shows a typical electron microscopy image of a Mo tip after coating with the a-D film. The a-D film exhibits a continuous and atomically smooth surface as expected for room temperature deposition. Such a morphology is entirely different from that of CVD diamond coatings, which are often rough and sometimes discontinuous. The SEM image also shows a few particles, 1000 Å in diameter, arising from melted droplets from the target, which usually occur in laser ablation.

A TEM image is shown in Fig. 2. The overall morphology of the a-D film shows a continuous and smooth surface. In addition, the coating conforms to the emitter shape. At the apex of the tip, however, the coating is denser than that at the



FIG. 2. TEM image of a Mo tip coated with a-D film, showing a nano-columnar microstructure. Note that the sizes and the angles of the columns vary from the tip to the shank.

shank. One of the most striking features of the TEM image is that the film shows a clear nano-columnar microstructure, which will be discussed below.

The areas of the apex and the shank were also examined by selected area electron diffraction. The SAED patterns confirmed that the films in both these areas are amorphous. Figure 3 is a typical SAED pattern from the tip; the diffuse rings show the amorphous nature of the film, while the bright spots represent single crystalline structure from the textured Mo substrate.

C. EELS measurements

Typical EELS spectra from a-D film coated Mo emitters are shown in Fig. 4. The peak centered at 286 eV results from excitations of electrons in the ground states $1s$ core levels to the vacant π^* state, while excitations to the higher lying σ^* states occur in the region of 290–305 eV. The atomic fraction of sp^3 bonded carbon was calculated to be around 50% at the apex and 30% at the shank.

D. Field emission characteristics

Figure 5 shows two typical current-voltage curves from one of the Mo tip emitters before and after a-D film coating. Prior to having the I-V measurements, the tips were conditioned for a few hours. A Fowler–Nordheim plot from the I-V data, as shown in Fig. 6, exhibits a straight line, indicating an electron tunneling process. Compared to the uncoated

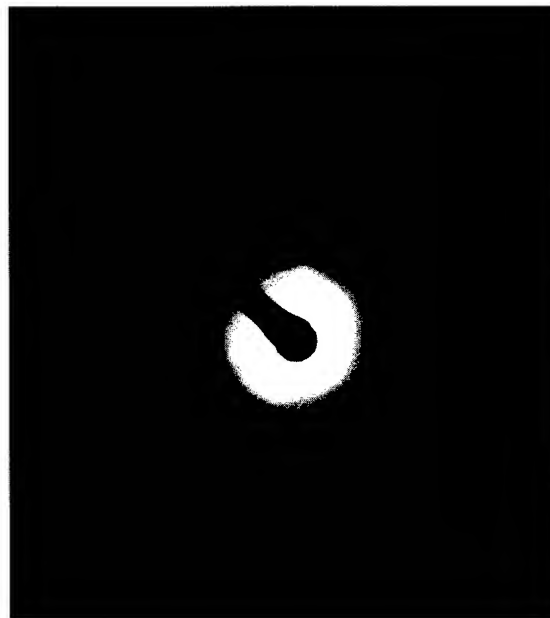


FIG. 3. SAED pattern of the Mo tip coated with a-D film, showing amorphous nature of the film. The bright spots are from the textured Mo substrate.

Mo tip, the coated emitter shows a substantial improvement in field emission performance with a lower turn-on voltage and a larger emitting current.

IV. DISCUSSIONS

A. Nano-columnar structure

From the TEM image, the a-D films show an interesting nano-columnar pattern, which reveals several features. First, the angle β of the columns to the surface normal varies from about 50° at the shank to 0° at the apex. Second, the size of the columns varies from about 200 Å at the shank to a few tens of Å at the apex. Third, the density of the columns in-

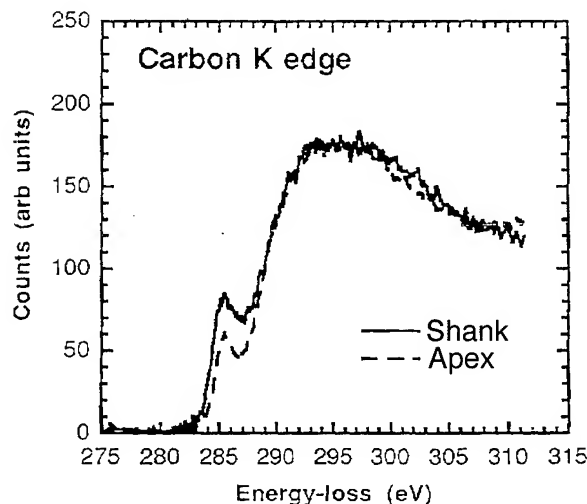


FIG. 4. EELS spectra from a-D films at the apex and the shank.

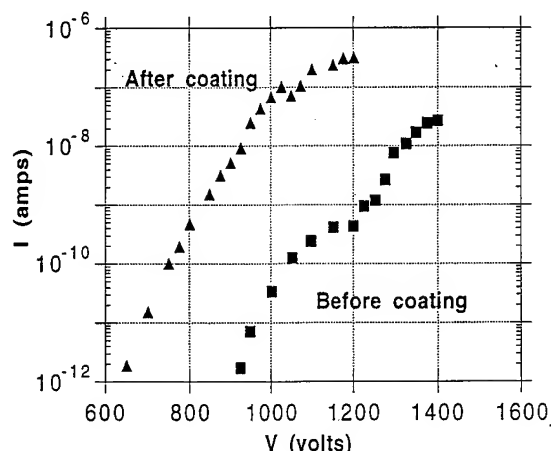


FIG. 5. Field emission current-voltage characteristics of the Mo tip emitter before and after deposition of a-D film.

creases from the shank to the apex. During the deposition, the incident particles (ions and atoms) are normal to the apex of the tip and tilted towards the curvature of the tip; therefore, the incident angle α of the particles to the surface normal varies from zero at the apex to about 65° at the shank. Calculations indicate that the incident angle α and the angle β roughly follow the so-called tangent rule:²⁵

$$\tan \alpha = 2 \tan \beta.$$

Such a nano-columnar structure may be explained by a shadowing effect and low mobility of adatoms during the deposition.^{25,26} As the shadowing effect decreases from the shank to the tip, the size of columns becomes smaller and there are fewer voids. Therefore, the film at the apex is denser than that at the shank. On the other hand, interactions between energetic carbon particles and surface atoms would play a role as well. Since the incident angle varies from 0° at the apex to 65° at the shank, one would expect an increased relaxation effect at the shank.²⁷ At a large incident angle, the

energetic ions and atoms tend to make more than one forward bouncing landing before finally being bonded to the surface atoms; for each of these landings, these ions and atoms would lose some of their energy. As a result of such a relaxation effect, these particles may not have adequate energy to form sp^3 bonding. However, at the apex, energetic particles would make a head-on impact on adatoms to form sp^3 bonding, or provide these in sp^2 bonding with some activation energy to surmount the barrier going from sp^2 to sp^3 , as proposed by Cuomo *et al.* in Ref. 27. This process would lead to a denser film with a higher fraction of sp^3 bonding, which in fact has been confirmed by EELS findings.

B. EELS data

Whereas the films at the apex and the shank show a distinct microstructure, the EELS data indicate a sp^3 content of 50% at the apex and 30% at the shank. However, these values are lower as compared to the values of 75–80% from planar samples. Such lower values most likely result from poor thermal conduction inherent to tip geometry. Energies of carbon ions and atoms during laser ablation can be as high as 80 and 12 eV, respectively, and increase with laser fluence.¹⁹ The deposition rate of the a-D film was calculated to be 0.16 \AA/pulse . For a pulse duration of 20 ns, the arrival rate of atoms and ions can amount to about 10^6 monolayers/s, among which 5–10% are ions.¹⁹ This is a very fast condensation process. Depending on energy distribution and the number of carbon ions and atoms, the tip with poor thermal conduction, could be heated up to a high temperature in the 20 ns period of time, resulting in reduced sp^3 bonds.²³ On the other hand, the even lower sp^3 content at the shank results from relaxation processes of incident particles ending up with the formation of more sp^2 bonded carbons, as discussed above. From the field emission point of view, a lower sp^3 content leading to a lower or zero-field emission from the shank is desirable, as what we need is a point emission right from the very tip end.

C. Field emission mechanisms

The I-V data from a-D film coated Mo tip emitter have shown a substantial improvement in field emission. Although the enhancement in emission from DLC/a-D film coated emitters (mostly planar ones) has been reported by several investigators,^{21–23} the mechanisms of field emission from DLC/a-D films are not available so far. However, some mechanisms proposed for field emission from diamond coated emitters may be applicable to a-D films. One is the field enhanced hot electron emission model for metal-insulator systems, developed earlier by Latham *et al.*²⁸ This model involves two processes: first, electrons tunnel from a metal into the conduction band of a dielectric through a potential barrier at the metal/dielectric interface. Second, the electrons gain energy from the electrical field, becoming “hot” as they pass through the dielectric, and finally escape into vacuum. From this model, field emission performance

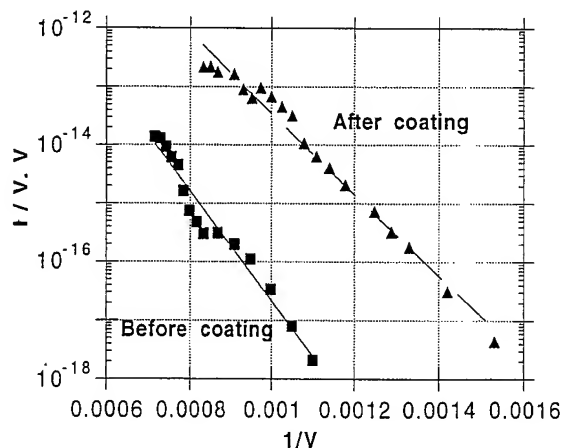


FIG. 6. Fowler-Nordheim plot of Mo tip emitter before and after deposition of a-D film.

depends on field penetration and the potential barrier at the interface. High carrier concentration in the conduction band of the dielectric would lead to a relatively thin depletion layer and hence a narrower potential barrier. For a-D films, an unpaired state density of $\geq 5 \times 10^{18} \text{ cm}^{-3}$ has been reported.²⁹ Although these defect levels lie well below mid-gap, the a-D film has a much narrower optical gap around 1.5–2.5 eV,¹⁷ as compared to the wide band gap (5.45 eV) of diamond. Therefore, it is possible for electrons to be excited into the conduction band through a hopping mechanism. Also, additional defects at the Mo/a-D interface are expected to contribute to a decrease in depletion layer width.

On the other hand, the nano-columnar structure of the a-D film may also play a role in field emission. For instance, the presence of nano-columns may lead to a distribution of the electric field within the a-D film and the interface. However, so far we do not know how exactly such a field distribution contributes to field emission. Another possibility is that the nano-columns at the tip serve as natural conduction channels. It was suggested³⁰ that graphite particles (sp^2 bonded carbons) contained in DLC/a-D films provide conduction channels for the emitted electrons, while nano-crystalline diamond particles (sp^3 bonded carbons) in DLC/a-D films have a negative electron affinity effect as seen in diamond films. The nano-columns at the apex contain about 50% sp^3 diamond carbon and 50% graphite carbon, which may play both these roles in field emission. Previous work²³ has indicated a-C films (mostly sp^2 bonds) would degrade field emission, therefore, there would be an optimal proportion of sp^3 content, which will provide best field emission. So far no one knows what the optimal sp^3 content of the film is. To answer this question, together with other questions such as the role of the nano-columns in field emission, more work needs to be done to correlate field emission with the bonding structure and micro-structure. We believe these studies would help to understand the mechanism of field emission from a-D films.

V. CONCLUSIONS

We have deposited a-D films on Mo tip emitters by pulsed laser ablation. HRTEM images of these films revealed a clear nano-columnar microstructure. At the shank, the density of the film was relatively low, while the film at the apex was much denser with smaller columns. EELS studies showed that the sp^3 content of the films was around 50% at the apex and 30% at the shank. SAED confirmed that the film was amorphous in nature. Field emission showed a substantial improvement with a reduced onset voltage and a higher current. We have demonstrated for the first time, to our knowledge, that a-D films can be deposited on Mo tip emitters by pulsed laser deposition to achieve a better field emission performance.

ACKNOWLEDGMENTS

The authors wish to express our thanks to Skion and Commonwealth Scientific for generous support for the

project. Dr. S. P. Bozeman is also thanked for helpful discussions.

- ¹C. Wang, A. Garcia, D. C. Ingramand, and M. E. Kordesh, *Electron Lett.* **27**, 1459 (1991).
- ²N. S. Xu, Y. Tzeng, and R. V. Latham, *J. Phys. D* **28**, 1776 (1993).
- ³J. Liu, V. V. Zhirnov, G. J. Wojak, A. F. Myers, W. B. Choi, J. J. Hren, S. D. Wolter, M. T. McClure, B. R. Stones, and J. T. Glass, *Appl. Phys. Lett.* **65**, 2842 (1994).
- ⁴E. I. Givargizov, V. V. Zhirnov, A. N. Stepanova, E. V. Rakova, A. N. Kiseley, and P. S. Plekhanov, *Appl. Surf. Sci.* **87/88**, 24 (1995).
- ⁵M. W. Geis, N. N. Efremow, J. D. Woodhouse, M. D. Mcaleese, M. Marchywka, D. G. Socker, and J. F. Hochedez, *IEEE Lett.* **12**, 456 (1991).
- ⁶K. OKano, K. Hoshina, M. Lida, S. Koizumi, and T. Inuzuka, *Appl. Phys. Lett.* **64**, 2742 (1994).
- ⁷W. B. Choi, J. J. Cuomo, V. V. Zhirnov, A. F. Myers, and J. J. Hren, *Appl. Phys. Lett.* **68**, 1 (1996).
- ⁸W. B. Choi, J. Liu, M. T. McClure, A. F. Myers, V. V. Zhirnov, J. J. Cuomo, and J. J. Hren, *J. Vac. Sci. Technol. B* **14**, 2050 (1996).
- ⁹F. J. Himpsel, J. A. Knapp, J. A. Van Vechten, and D. E. Eastman, *Phys. Rev. B* **20**, 624 (1979).
- ¹⁰J. Van der Weide, Z. Zhang, P. K. Baumann, M. G. Wensell, J. Bemhole, and R. J. Nemanich, *Phys. Rev. B* **50**, 5903 (1994).
- ¹¹D. R. McKenzie, D. Muller, B. A. Pailthorpe, Z. H. Wang, E. Kravtchinskaia, D. Segal, P. B. Lukins, P. D. Swift, P. J. Martin, G. Amaratunga, P. H. Gaskell, and A. Saeed, *Diam. Relat. Mater.* **1**, 51 (1991).
- ¹²C. B. Collins, F. Davanloo, E. M. Juengerman, W. R. Osborn, and D. R. Jander, *Appl. Phys. Lett.* **54**, 216 (1988).
- ¹³T. A. Friedmann, M. P. Siegal, D. R. Tallant, R. L. Simpson, and F. Dominguez, in *Novel Forms of Carbon II*, edited by C. L. Renschler, D. Cox, J. Pouch, and Y. Achiba (Materials Research Society, San Francisco, 1994), Vol. 1, p. 349.
- ¹⁴Y. Lifshitz, S. R. Kasi, and J. W. Rabalais, *Phys. Rev. Lett.* **62**, 1290 (1989).
- ¹⁵E. J. G. Spencer, P. H. Schmidt, D. C. Joy, and F. J. Sansalone, *Appl. Phys. Lett.* **29**, 118 (1976).
- ¹⁶J. J. Cuomo, D. L. Pappas, J. Bruley, J. P. Doyle, and K. L. Saenger, *J. Appl. Phys.* **70**, 1706 (1991).
- ¹⁷R. L. David, L. Pappas, R. A. Roy, J. J. Cuomo, and V. M. Sura, *Appl. Phys. Lett.* **61**, 171 (1992).
- ¹⁸T. A. Friedmann, K. F. McCarty, J. C. Barbour, M. P. Siegal, and D. C. Dibble, *Appl. Phys. Lett.* **68**, 1643 (1996).
- ¹⁹D. L. Pappas, K. L. Saenger, J. J. Cuomo, and R. W. Dreyfus, *J. Appl. Phys.* **72**, 3966 (1992).
- ²⁰C. Xie, N. Kumar, and C. B. Collins, *Proceedings of the 6th IVMC*, 12–15 July, 1993, Newport, RI, p. 162.
- ²¹G. A. J. Amaratunga and S. R. P. Silva, *Appl. Phys. Lett.* **68**, 2529 (1996).
- ²²F. Y. Chuang, C. Y. Sun, H. F. Cheng, C. M. Huang, and I. N. Lin, *Appl. Phys. Lett.* **68**, 1666 (1996).
- ²³A. F. Myers, S. M. Camphausen, J. J. Cuomo, and J. J. Hren, *J. Vac. Sci. Technol. B* **14**, 2024 (1996).
- ²⁴J. Y. Shim, E. J. Chi, and H. K. Baik, *Proceedings of the 9th IVMC*, 7–12 July, 1996, St. Petersburg, Russia, p. 263.
- ²⁵J. M. Nieuwenhuizen and H. B. Haanstra, *Philips Tech. Rev.* **27**, 87 (1966).
- ²⁶Milton Ohring, *The Materials Science of Thin Films* (Academic, New York, 1992).
- ²⁷J. J. Cuomo, D. L. Pappas, R. Lossy, J. P. Doyle, J. Bruley, G. W. Di Bello, and W. Krakow, *J. Vac. Sci. Technol. A* **10**, 3414 (1992).
- ²⁸K. H. Bayliss and R. V. Latham, *Proc. R. Soc. London, Ser. A* **403**, 285 (1986).
- ²⁹G. A. J. Amaratunga, J. Robertson, V. S. Veerasamy, W. I. Milne, and D. R. McKenzie, *Diam. Relat. Mater.* **4**, 637 (1995).
- ³⁰Z. Feng, I. G. Brown, and J. W. Ager, *J. Mater. Res.* **10**, 1585 (1995).

Oxygen implantation induced interdiffusion in AlGaAs/GaAs quantum well structures

Padraig J. Hughes and Bernard L. Weiss^{a)}

*School of Electronic Engineering, Information Technology and Mathematics, University of Surrey,
Guildford, Surrey GU2 5XH, United Kingdom*

Spirit Tlali and Howard E. Jackson

Department of Physics, University of Cincinnati, Cincinnati, Ohio 45220-0011

(Received 20 September 1996; accepted 11 April 1997)

The effects of low dose (10^{14} ions/cm²) oxygen implantation on the subband structure of an AlGaAs/GaAs single quantum well is reported here using photoluminescence and photoreflectance spectroscopy. Postimplantation rapid thermal annealing at 900 °C up to 180 s induces compositional interdiffusion resulting in enhanced transition energy shifts and a reduction in the photoluminescence count rate. Comparisons of transition energy shifts show that the electron two to heavy-hole two transition energy is more sensitive than the ground state electron to heavy-hole transition for these annealing conditions. © 1997 American Vacuum Society. [S0734-211X(97)03504-X]

I. INTRODUCTION

Thermally induced interdiffusion is an important post-growth means of modifying the optical properties of quantum well (QW) structures.¹ The rate of interdiffusion, which is determined by the interdiffusion coefficient, may be enhanced by the presence of either vacancies² or impurities.³ In particular, the implantation of impurities is well established due to the high degree of spatial and depth control of the impurity concentration.⁴ More importantly, it is known that the implantation of selected ions enhance the interdiffusion rate, thereby demonstrating its ability to control the QW optical properties for device applications. Typical applications include wavelength division multiplexing, where an array of lasers is required each with slightly different operating wavelengths,⁵ the fabrication of low loss waveguides,^{6,7} and devices for optical integrated circuits.

Since most device fabrication processes include high temperature processing, it is important to characterize the effects of thermally induced QW interdiffusion. Interdiffusion is characterized optically by monitoring the behavior of the QW interband transitions as a function of the annealing parameters. The most well established optical technique to characterize interdiffused QWs is low temperature photoluminescence (PL) which determines only the ground state transition energies.⁸ However, since the higher order transitions also contribute to the QW optical properties, other techniques, such as photoluminescence excitation spectroscopy or photoreflectance (PR), are required to characterize the higher order transitions and therefore completely evaluate QW interdiffusion.

This article discusses the effects of oxygen implantation as a function of annealing parameters on the ground state electron to heavy-hole transition energy (H_{11}) and the electron two to heavy-hole two transition energy (H_{22}) of an AlGaAs/GaAs single QW structure using PL and PR, respec-

tively. Oxygen ions were chosen for two reasons. First, oxygen implantation into GaAs⁹ and AlGaAs¹⁰ is well known for the removal of free carriers and the electrical isolation of semiconductor devices.¹¹ Second, it has been demonstrated previously that the presence of implanted oxygen enhances the rate of interdiffusion rates in both AlGaAs/GaAs¹² and InAlAs/InGaAs QW structures.¹³

II. EXPERIMENTAL METHOD

The QW structure was grown by metalorganic vapor phase epitaxy (MOVPE) and had a single 100-Å-thick GaAs well layer, 1000 Å Al_{0.3}Ga_{0.7}As barrier layers, and was terminated with a 50 Å cap. The implants were carried out at room temperature using an oxygen ion dose of 1×10^{14} cm⁻² and an ion energy of 75 keV was used to give a projected ion range R_p which corresponds to the center of the well layer, i.e., $R_p \pm \Delta R_p$ of 1140 ± 488 Å.¹⁴ After implantation, the samples were annealed to remove the implantation induced lattice damage and to promote compositional interdiffusion. Two sets of samples, i.e., as-grown and oxygen implanted, were annealed in a nitrogen rich atmosphere at 900 °C for annealing times of 10, 60, 120, and 180 s using a GaAs proximity cap. PL measurements were performed using a 100 mW Ar⁺ ion laser ($\lambda = 514.5$ nm). The samples were placed in a cryostat immersed in liquid helium ($T = 4$ K) and the spectra were collected and analyzed using a 1 m spectrometer. PR measurements were carried out using a conventional room temperature PR system which has been described elsewhere.¹⁵

III. RESULTS

A. As-grown samples

The PL and PR spectra of the as-grown QW structure are shown in Figs. 1(a) and 2(a), respectively. The two peaks in the PL spectrum at 1.546 eV and 1.495 eV were identified as the H_{11} and that due to residual carbon doping arising from

^{a)}Electronic mail: B.Weiss@ee.surrey.ac.uk

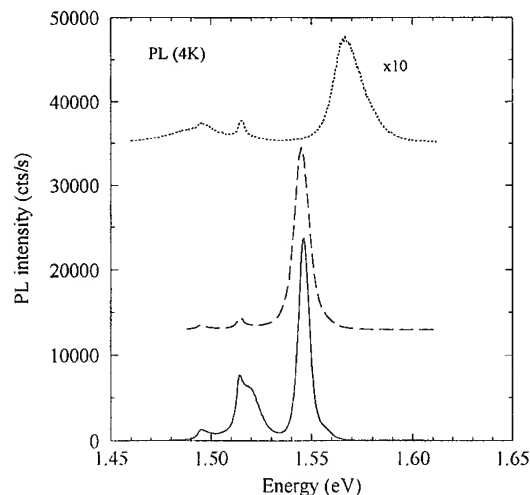


FIG. 1. PL spectra of (a) the as-grown (—), (b) annealed (---), and (c) oxygen implanted/annealed (···) QW structure at 900 °C for 120 s.

the MOVPE growth, respectively. The background carbon has been identified previously using phase sensitive PR.¹⁶ The peak at 1.515 eV, which has a broad high energy tail, is attributed to a neutral donor/acceptor bound exciton transition.¹⁷ The PR spectrum shows strong Franz-Keldysh oscillations (FKO) above the GaAs band edge were attributed to a built-in electric-in field of ~ 30 kV/cm in the structure at the QW substrate interface.¹⁵ The FKO were found to mask the H_{11} PR spectral features which made its identification impossible without detailed modeling. However, the H_{22} at 1.541 eV¹⁸ was determined by fitting the PR spectrum using Aspnes third derivative functional forms.¹⁹ Theoretical transition energies were calculated and compared to experimental values determined using different QW profiles.^{18,20} Characterization of this structure has shown that the well barrier interface is graded rather than abrupt and fitting has shown that it is best described by an exponential function.^{18,20}

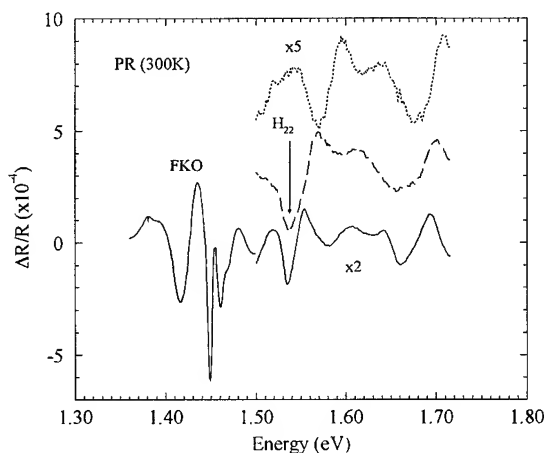


FIG. 2. PR spectra of the (a) as-grown (—), (b) as-grown annealed (---), and (c) oxygen implanted and annealed (···) QW structure at 900 °C for 120 s.

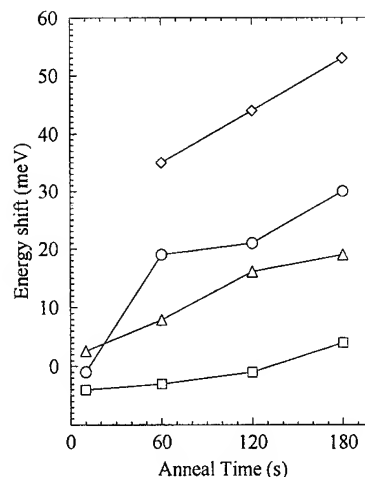


FIG. 3. The variation of the experimentally measured change in both the H_{11} (□) and H_{22} (Δ) energies for the as-grown and the H_{11} (○) and H_{22} (◇) energies for the implanted samples as a function of anneal time. [Note: It was not possible to determine the H_{22} (◇) energy in the implanted sample after annealing for 10 s from the PR spectrum.]

The effect of annealing at 900 °C for 120 s on the PL and PR spectra of the as-grown sample, which are shown in Figs. 1(b) and 2(b), respectively, demonstrate that it has no significant effect on H_{11} and H_{22} of the as-grown sample. A recent study of the thermal stability of this structure and other similar structures at 800, 900, and 1000 °C resulted in limited interdiffusion and produced an interdiffusion coefficient of $2.9 \pm 0.4 \times 10^{-17}$ cm²/s at 900 °C and an activation energy of 3.3 ± 0.5 eV.²¹ For example, annealing at 900 °C for 120 s results in an interdiffusion length L_d of only ~ 6 Å.²¹

B. Implanted samples

Significant shifts of both H_{11} and H_{22} are seen in the PL and PR spectra of the implanted samples, respectively, see Figs. 1(c) and 2(c). The change in energy for both transitions as a function of anneal time are shown in Fig. 3, where it is assumed that the energy shifts are independent of the sample temperature during measurement. These results show that H_{22} is more sensitive to annealing than H_{11} for the annealing conditions used here. This result agrees with theoretical calculations for transition energy shifts as a function of interdiffusion length L_d (which is proportional to the square root of the anneal time),²² which are shown in Fig. 4. L_d has been normalized to the as-grown well widths L_z in Fig. 4 in order to compare accurately the effect of interdiffusion for different L_z . The well widths modeled were 100 and 125 Å since previous PR analysis of this wafer has shown that L_z varies between 110 and 120 Å, depending upon which interband transition is used for the fitting.^{18,20} The results of Figs. 3 and 4 are explained in terms of changes in both the confinement energy and the compositional dependence of the QW band gap energy.²² These results show that the higher order transition energy shifts are more sensitive than their ground state counterparts for $L_d/L_z < 0.2$,²² i.e., during the early stages of annealing. The reduction of H_{11} between the as-grown

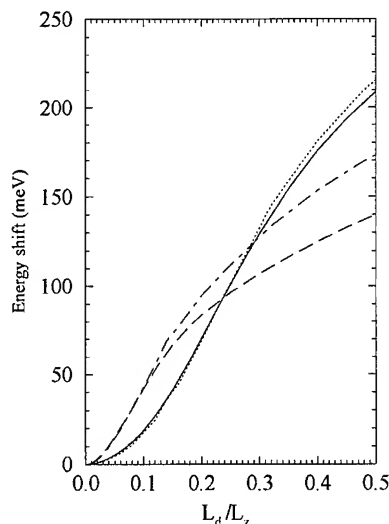


FIG. 4. The theoretical variation of the H_{11} and H_{22} energies, respectively, with L_d/L_z for well widths of 100 Å (—, ---) and 125 Å (···, -·-·-).

sample and after annealing for 10 s, as shown in Fig. 3, is not a real effect. It is attributed to variations of the QW parameters (e.g., well thickness and barrier composition) across the wafer, as shown by PL measurements of two as-grown samples from the same wafer which show a 2 meV difference in H_{11} .

The implanted samples show an increase in both H_{11} and H_{22} which is attributed to the presence of the oxygen in the samples, see Fig. 3. It is clear that the change of H_{22} in the implanted samples after annealing is approximately double that compared to the as-grown samples for the same annealing cycle. This is reflected in the value of $L_d \sim 12$ Å for the implanted sample after annealing at 900 °C for 120 s. Although annealing at 120 s reduces the extent of the implantation induced damage in the sample, it is not clear whether the enhanced interdiffusion observed here is due to the annealing of the damage or an oxygen related defect mechanism or both and further work is required here.

The PL emission count rate of the H_{11} and the GaAs band edge related signals were determined using constant excitation conditions for both the as-grown and implanted samples with annealing time, see Fig. 5. The count rate, which is a

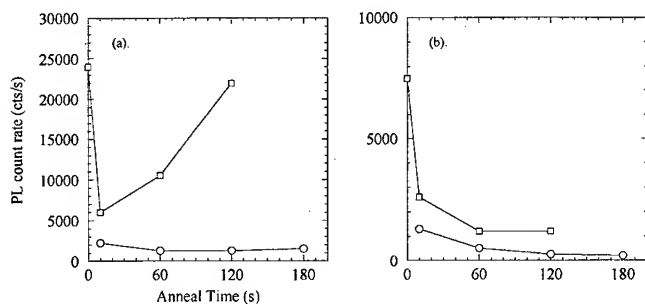


FIG. 5. The variation of the PL count rate with anneal time for the as-grown (\square) and implanted (\circ) samples for (a) H_{11} and (b) GaAs band edge signals.

measure of the free carrier concentration in the sample, drops dramatically during the first 10 s anneal for both the as-grown and implanted samples. For the as-grown sample, the H_{11} count rate recovers to a value similar to the as-grown value after annealing for 120 s. This initial drop is attributed to defects induced by annealing during the first 10 s arising from the poor performance of the proximity cap. In contrast, the GaAs band edge related emission remains quenched after 180 s, i.e., these GaAs related defects have been partially annealed out.¹⁷

For the oxygen implanted samples, both the H_{11} and band edge emission count rate are low and approximately constant for all anneal times. The low count rate is attributed to the trapping of the free carriers by the implanted oxygen whose ion energy was calculated so that the peak oxygen concentration occurred at the center of the GaAs well layer. This oxygen induced reduction of the PL signal has been reported previously for metalorganic chemical vapor deposition grown AlGaAs²³ and has been attributed to the inclusion of oxygen. The lack of recovery of the PL count rate after annealing is attributed to the stability of the oxygen induced free carrier removal mechanism which demonstrates that this process may be used for both band gap modifications and carrier removal in the QW structure. These effects are important for the electrical and optical isolation of devices²⁴ on a single substrate, for the optimization of device characteristics, and for the fabrication of optical integrated circuits.

IV. CONCLUSIONS

The results presented show the differences in the variation of the QW transition energies during the early stages of interdiffusion. The presence of implanted oxygen is shown to more than double the rate of interdiffusion in AlGaAs/GaAs single QW structures. The results also show that H_{22} is more sensitive to annealing than H_{11} , as has been predicted previously. Since the optical properties are determined by all transitions in the QW structure, this result demonstrates that it is essential to determine the behavior of all QW transitions to predict the optical properties of the QW structure.

The effect of a low dose oxygen implant into a single QW structure is shown to promote interdiffusion in addition to providing a stable strong carrier removal process. This result shows that it is possible to use a single process to simultaneously induce two effects which selectively modify the optical and electrical properties of the structure, i.e., the implantation of oxygen, which enhances interdiffusion and subsequently modifies the optical properties of the structure, while the presence of oxygen removes free carriers from the structure.

ACKNOWLEDGMENT

The authors wish to acknowledge the Dr. J. Roberts of the University of Sheffield for growth of the QW structures.

¹D. G. Deppe and N. Kolonyak, J. Appl. Phys. **64**, R93 (1988).

²M. Ghisoni, A. W. Rivers, K. Lee, G. Parry, X. Zhang, A. Stanton-Bevan, M. Pate, G. Hill, C. Button, and J. S. Roberts, Mater. Res. Soc. Symp. Proc. **262**, 823 (1992).

- ³I. Harrison, *J. Mater. Sci.* **4**, 1 (1993).
- ⁴B. L. Weiss, I. V. Bradley, N. J. Whitehead, and J. S. Roberts, *J. Appl. Phys.* **71**, 5715 (1992).
- ⁵R. D. Burnham and R. L. Thornton, *Inst. Phys. Conf. Series No.* **83**, 9 (1986).
- ⁶Y. Suzuki, H. Iwamura, and O. Mikami, *IEEE Photonics Technol. Lett.* **2**, 818 (1990).
- ⁷M. Kumar, V. Gupta, G. De Brabander, P. Chen, J. T. Boyd, A. J. Steckl, A. G. Choo, H. E. Jackson, R. D. Burnham, and S. C. Smith, *IEEE Photonics Technol. Lett.* **4**, 435 (1993).
- ⁸J. D. Ralston, S. O'Brien, G. Wicks, and L. F. Eastman, *Appl. Phys. Lett.* **52**, 1511 (1988).
- ⁹N. J. Whitehead and B. J. Sealy, *Electron. Lett.* **26**, 866 (1990).
- ¹⁰S. J. Pearton, M. P. Iannuzzi, C. L. Reynolds, and L. Peticolas, *Appl. Phys. Lett.* **52**, 395 (1988).
- ¹¹S. J. Pearton, *Mater. Sci. Rep.* **4**, 313 (1990).
- ¹²F. Xiong, T. A. Tombrello, C. L. Schwartz, and S. A. Schwartz, *Appl. Phys. Lett.* **57**, 896 (1990).
- ¹³E. K. Rao, P. Ossart, H. Thibierge, M. Quillec, and P. Krauz, *Appl. Phys. Lett.* **57**, 2190 (1990).
- ¹⁴I. R. Chakarov, S. Todorov, and D. S. Karpuzov, *Nucl. Instrum. Methods Phys. Res. B* **69**, 193 (1992).
- ¹⁵P. J. Hughes, B. L. Weiss, and T. J. C. Hosea, *J. Appl. Phys.* **77**, 6472 (1995).
- ¹⁶P. J. Hughes and B. L. Weiss, *J. Vac. Sci. Technol. B* **14**, 632 (1996).
- ¹⁷T. Hiramoto, Y. Mochizuki, T. Saito, and T. Ikoma, *Jpn. J. Appl. Phys.* **24**, L921 (1985).
- ¹⁸P. J. Hughes, T. J. C. Hosea, and B. L. Weiss, *Semicond. Sci. Technol.* **9**, 1339 (1995).
- ¹⁹D. E. Aspnes *Surf. Sci.* **37**, 418 (1973).
- ²⁰W. C. H. Choy, P. J. Hughes, and B. L. Weiss, *Mater. Res. Soc. Symp. Proc.* **450**, 425 (1997).
- ²¹P. J. Hughes, E. H. Li, and B. L. Weiss, *J. Vac. Sci. Technol. B* **13**, 2276 (1995).
- ²²P. J. Hughes, B. L. Weiss, and H. E. Jackson, *Semicond. Sci. Technol.* **12**, 808 (1997).
- ²³G. B. Stringfellow and G. Hom, *Appl. Phys. Lett.* **34**, 794 (1979).
- ²⁴K. L. Tsai, C. P. Lee, K. H. Chang, H. R. Chen, and J. S. Tang, *J. Appl. Phys.* **76**, 274 (1994).

Observation of hot electron relaxation in GaAs/AlGaAs multiple quantum wells by excitation spectroscopy*

H. Z. Wu,^{a)} J. H. Liu, G. O. Dong, and J. Z. Wu
Department of Physics, Hangzhou University, Zhejiang 310028, China

Z. Z. Ye and X. B. Jiang
State Key Laboratory of Silicon Material Science, Zhejiang University, Hangzhou 310027, China

(Received 18 September 1996; accepted 23 May 1997)

Hot electron relaxation in GaAs/AlGaAs multiple quantum well (MQW) structure was studied with the use of photoluminescence excitation (PLE) spectroscopy. Oscillation due to the emission of confined longitudinal optical (LO) GaAs phonons, by photoexcited electrons were observed in the (6 K) excitation spectra. The period of the oscillation is different from that observed in bulk GaAs. The calculation from a four-band Kane model, describing the mixing of heavy- and light-hole bands at wave vector away from $k=0$, was used to interpret the difference of oscillation features between the GaAs/AlGaAs MQW structure and bulk GaAs. The recorded PLE spectrum and calculated results show that photoexcited electrons can directly cascade downwards to the exciton energy state by LO phonon emissions. © 1997 American Vacuum Society. [S0734-211X(97)14704-7]

I. INTRODUCTION

Hot electron relaxation in low-dimensional semiconductor structures has attracted significant interest.¹⁻³ This is not only because of its intrinsic effects to semiconductors, but also because of the realities in many devices,^{4,5} such as quantum well lasers, detectors, and modulation doped field transistors, etc. Optical phonon scattering in two-, one-, and zero-dimensional structures has been observed in photoluminescence (PL) and photoluminescence excitation (PLE). The "phonon bottleneck" effect in GaAs/AlGaAs quantum well dots⁶ and the exciton-phonon interaction due to strong lateral localization of the excitons in the GaAs channels⁷ were investigated using PL and PLE techniques. It is known that hot electrons with energy higher than the band gap of a semiconductor can relax down to emitting states by the emission of phonons. When the excess energy of photoexcited electrons is much larger than the energy of longitudinal optical (LO) phonons (36.6 meV for three-dimensional GaAs), hot electrons interact with LO phonons strongly; such that the electrons preferentially lose excess energy through the emissions of LO phonons. When the excess energy of hot electrons is smaller than the energy of a LO phonon it is impossible to emit LO phonons, and the hot electrons relax excess energy by the emission of acoustic phonons. However, the scattering rate of acoustic phonon emissions is several orders of magnitude lower than the LO phonon scattering rate. This will raise the probability of the hot electron's escape into the outermost GaAs, rather than forming an exciton. Hence, in the process of electron-acoustic phonon scattering, a lower luminescence efficiency could be accompanied that provides an effective way to observe different energy relaxation processes of hot electrons. Such an expectation was first observed in PLE spectra by Weisbuch⁸ from three-dimensional GaAs samples, which exhibited a series of oscillations attrib-

utable to LO phonon emissions by relaxing hot electrons. The oscillation period was almost constant. Similar results were also observed via oscillations in the spectral responses of photoconductivity (PC) from bulk GaAs.⁹ The energy period in either PLE or PC for a system with parabolic conduction and valence bands is given by $\hbar\omega(1+m_e/m_h)$, here $\hbar\omega$ is the LO phonon energy, m_e is the electron effective mass, and m_h could be either heavy- or light-hole effective mass depending on the origin of the electron-hole pair created by the photoexcitation. The enhancement factor of $(1+m_e/m_h)$ reflects the creation of electron-hole pairs away from $k=0$. Similar oscillatory structure was observed in the 4 K PLE spectrum from aGaAs/Al_{0.3}Ga_{0.7}As superlattice, where the binary and alloy components were both only 10-Å-thick.¹⁰ In this article, a high quality GaAs/AlGaAs multiple quantum well (MQW) structure was studied using the PLE technique. We report the first observation of LO phonon emissions in the process of hot electron relaxation to emitting states from the MQW. The recorded oscillation in the PLE spectrum is different from that of bulk GaAs as described above. This difference is attributed to the nonparabolicity and warping of subbands in the plane of the quantum wells. The nonparabolicity arises from the lift of the lowest subband energy somewhat above the bottom of the GaAs conduction band. The warping of subbands is due to the mixing effect of heavy- and light-hole bands at $k \neq 0$. To interpret the complicated PLE spectral response, here a four-band Kane model including heavy- and light-hole band mixing was used to calculate the energy dispersion of the subbands.

Before proceeding to a description of the experimental results and discussion, we should bear in mind the optical transition physics involved in a PLE spectrum. PLE has proved a particularly useful technique for studying the physics in low dimensional semiconductor structures, especially for the measurement of various subband gap energies in a whole range of quantum well and superlattice systems.¹¹⁻¹³

*No proof corrections received from author prior to publication.

^{a)}Electronic mail: hzwu@whalo.hzuv.edu.cn

The technique involves monitoring the PL intensity from a structure, while scanning the energy of excitation source. As the strength of the absorption coefficient increases with increasing excitation energy above the fundamental gap this leads to an increase of PL intensity. However, PLE differs in one major aspect to absorption spectroscopy, that the experimental process involves the relaxation of the photoexcited carriers down to the emitting states. Therefore, it is expected that a PLE spectrum contains information on the energy relaxation processes of hot electrons.

II. EXPERIMENTS

The multiple quantum well structure studied in this work was grown in a Varian Gen II molecular beam epitaxy system. The layers were deposited on (001) oriented semi-insulating GaAs substrates at a substrate temperature of 630 °C. The growth sequence was as follows: (1) 1.0 μm of GaAs buffer layer; (2) 0.13 μm AlGaAs cladding layer; (3) 60 periods of GaAs wells and AlGaAs barriers, with well and barrier widths of 75 and 170 Å, respectively; and (4) 0.13 μm cladding layer. Al fraction in AlGaAs barriers and cladding layers was 0.35. The barriers were thick enough to ensure that the GaAs wells were electrically decoupled. The quantum well thicknesses were measured by optical and x-ray diffraction techniques.

The PL and PLE spectra were measured at 6 K with the sample mounted on the cold finger of a variable temperature continuous flow cryostat. The sample was excited with light from an argon-ion-pumped tunable dye laser at a power density of $<1 \text{ W cm}^{-2}$. The PL was collected from the front surface of the sample and energy dispersed by a 1 m single grating spectrometer and detected by a cooled GaAs photomultiplier, and associated lock-in amplification techniques. PLE measurements were made by monitoring PL intensity at appropriate luminescent wavelength, while scanning the energy of the excitation source. Thus, hot electron relaxation processes could be expected to be recorded in the PLE spectrum.

III. RESULTS AND DISCUSSION

Figure 1 shows the PL spectrum (6 K) from the multiple quantum well structure and Fig. 2 shows the spectral response of the PLE at 6 K. The detected energy position in Fig. 2(a) was set in the low-energy tail of the $1s n=1$ electron-to-heavy-hole ($e-hh$) exciton state at 1.563 eV, and the detected energy position in Fig. 2(b) was set as $2s n=1$ ($e-hh$) exciton state at 1.581 eV. From both Figs. 1 and 2, we can see the sample studied is of high quality from an optical point of view; in that the spectra show sharp exciton lines with very narrow full width at half-maximum (FWHM), although the structure consists of 60 quantum wells. By comparing Figs. 1 and 2, the dominant peak at 1.571 eV in Fig. 1 obviously corresponds to a $1s n=1$ ($e-hh$) exciton line, that is confirmed by the subband structure calculation described in the following paragraphs. The FWHM of the peak is only 3.4 meV. However, in the PLE spectrum shown in Fig. 2(a), the sharp peak is resolved into

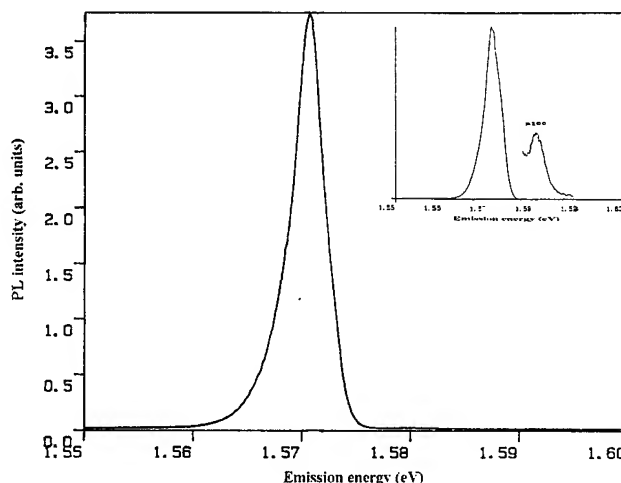
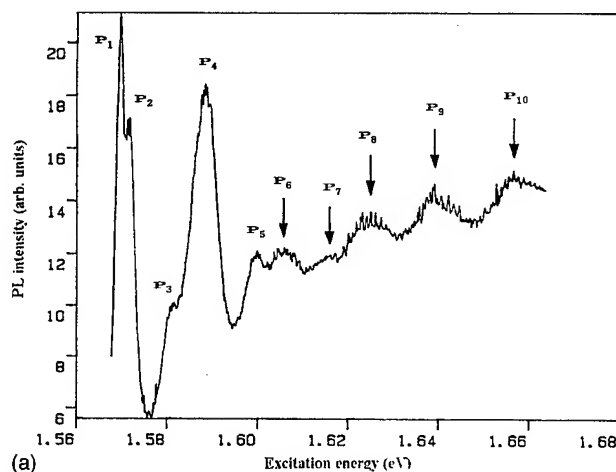
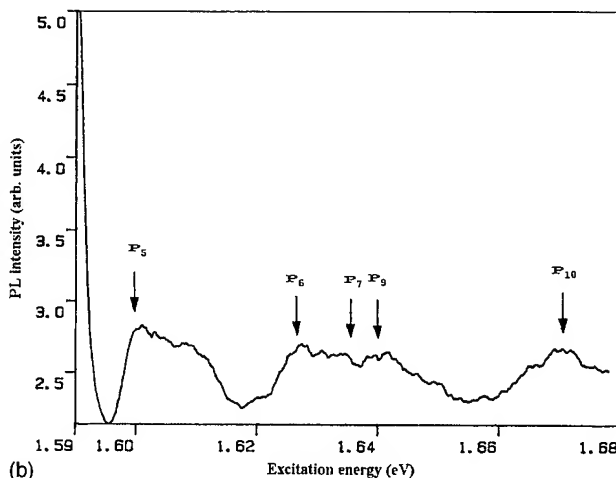


FIG. 1. The 6 K photoluminescence spectrum measured from the GaAs/Al_{0.35}Ga_{0.65}As MQW. The inset at the upper-right-hand corner shows the magnified result of the additional feature on the high-energy side of the $1s n=1$ ($e-hh$) exciton line.



(a)



(b)

FIG. 2. The 6 K photoluminescence excitation spectra measured from the GaAs/Al_{0.35}Ga_{0.65}As MQW and the arrows showing the measured peak energy positions. (a) Detected photon energy at 1.563 eV and (b) detected photon energy at 1.581 eV.

two close peaks positioned at 1.5700 eV (P_1) and 1.5715 eV (P_2). Our further temperature scanning of PLE measurements confirmed that P_1 and P_2 are $1sn=1$ ($e-hh$) bound exciton and free exciton lines, respectively.

From Fig. 1 we can see the clearly resolved additional feature on the high-energy side of the $1sn=1$ ($e-hh$) exciton line. The inset in Fig. 1 shows the magnified result of the smaller peak. Its energy position is at 1.581 eV, which is also clearly resolved in the PLE spectrum as P_3 . This additional peak was identified as the $2sn=1$ ($e-hh$) exciton transition. A similar peak was previously observed and was attributed to $2sn=1$ ($e-hh$) exciton recombination.^{11,14} The energy difference between the $2s$ line and the $1sn=1$ ($e-hh$) line is 12 meV, while is smaller than the LO phonon energy of bulk GaAs. We believe that the relatively strong transition of the $2s$ exciton line is due to the fact that the only available energy loss mechanism is acoustic phonon emission, since the amount of energy to be lost by free carriers close to the bottom of conduction is less than the LO phonon energy. The other most prominent peak at 1.589 eV (P_4) in Fig. 2(a) is evidently the $1s$ state of the $n=1$ light-hole (lh) exciton. Similarly, the peak at 1.600 eV (P_5) possibly represents the $2s$ replica of the $n-1$ ($e-lh$) exciton.¹⁴

Now we concentrate on the oscillatory features at the higher energy region shown in Fig. 2(a). The peaks in the higher energy region are at ~ 1.606 eV (P_6), 1.616 eV (P_7), 1.625 eV (P_8), 1.639 eV (P_9), and 1.656 eV (P_{10}). Obviously, the oscillation does not show a constant period, which is different from the oscillation feature of bulk GaAs.⁸ The optical transitions at higher energy region refer to absorption and recombination processes away from $k=0$. In the case of quantum wells, the lowest subband energy is somewhere above the bottom of the GaAs conduction-band edge. So, the nonparabolicity of the subbands away from $k=0$ needs to be considered, when interpreting the optical transitions related to these higher energy peaks. Furthermore, the mixing effect of heavy- and light-hole bands also contributes to the warping of the hole bands away from $k=0$. These effects make it complicated to interpret the oscillation feature shown in Fig. 2. To understand the oscillation feature, we need the knowledge of energy dispersion of subbands. Theoretical work on the band structures of GaAs/AlGaAs quantum wells was beautifully done by several groups.^{15–20} Here a four-band Kane model developed by Eppenga *et al.*²⁰ was used to calculate the energy dispersion that includes the nonparabolicity and warping of subbands of the MQW structure. The model combined a unified description of electron and hole states, with only five adjustable parameters for each material constituting of the MQW. It fully accounted for the coupling between the lowest electron, the heavy-hole, the light-hole, and the spin-orbit split-off hole bands and the coupling to all other bands was taken into account perturbatively. The calculated energy dispersion of band energies is shown in Fig. 3. From Fig. 3 we can see that the subband structure, especially heavy- and light-hole bands, is different from that of the three-dimensional GaAs. Hence, it is expected that hot

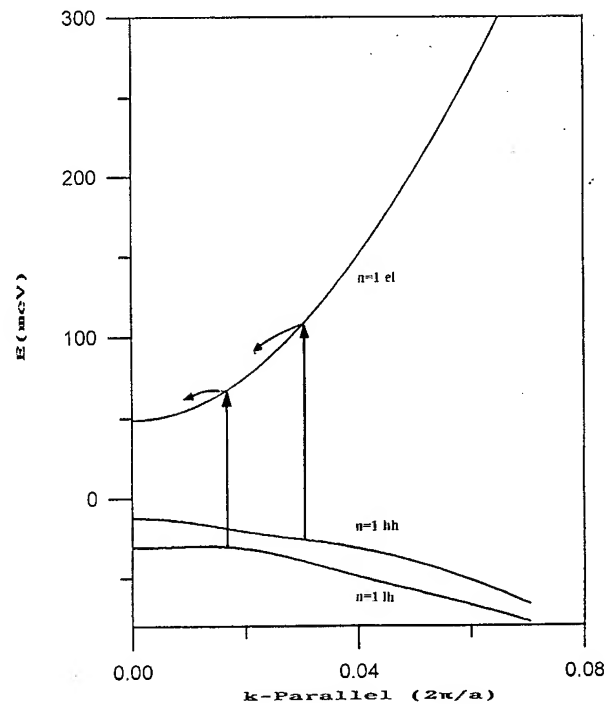


Fig. 3. The calculated energy dispersions of $n=1$ el , hh , lh subbands in [001] direction from the GaAs/Al_{0.35}Ga_{0.65}As MQW and the arrows showing the optical transitions. The band gap of the three-dimensional GaAs is not given in the energy axis.

electron relaxation processes in the GaAs/AlGaAs MQW structure will be different from that of the bulk GaAs.

Here Fig. 3 is used to interpret the optical transitions at the higher energy region shown in Fig. 2(a); because PLE as opposed to absorption reflects, both the creation of electron-hole pairs and their subsequent relaxation down to the emitting state, before proceeding to further discussion we should keep in mind the facts that have effects on the PLE spectrum: (a) the absorption coefficient of the quantum states, that is dependent on density of states; (b) the competition between relaxation and recombination of hot carriers in the path of cascade down to low energy states; (c) the detected photon energy. The peak P_9 at 1.639 eV shown in Fig. 2(a) was presumably attributed to the $n=1$ electron band to $n=3$ heavy-hole band transition at $k=0$. This is because the calculated $1e-3hh$ band gap was almost consistent with the energy of P_9 , while other quantum state transitions are either physically prohibited or at energies far from P_9 position. Also, this attribution was supported by the further PLE measurements, that as the detection energy was changed the positions of P_1 , P_2 , P_3 , P_4 , P_5 , and P_9 remained unchanged, while the positions of P_6 , P_7 , P_8 , and P_{10} would shift. From the PLE spectrum shown in Fig. 2(b) in which the detected energy was set at $2s$ $n=1$ ($e-hh$) exciton line (1.581 eV), we can see that P_5 and P_9 almost remained at same positions, while the P_6 (1.606 eV), P_7 (1.616 eV), and P_{10} (1.656 eV) in Fig. 2(a) moved to new positions at ~ 1.625 , 1.635, and 1.672 eV in Fig. 2(b), respectively. Here we attribute the P_6 , P_7 , P_8 , and P_{10} to the results of PL phonon

scattering, while hot electrons are relaxing to lower energy states rather than from inter subband transitions at $k=0$. As discussed in the following: this assignment could be supported by the results calculated from the band energy dispersion as shown in Fig. 3.

Figure 2(a) shows that the first peak (P_6) of the oscillation is at ~ 1.606 eV. However, from the calculation of subbands of the MQW structure shown in Fig. 3, we can see the bottom of the lowest conduction band ($n=1e$) is at 1.578 eV. The energy difference between P_6 and the bottom of the $n=1el$ band is only 28 meV, much smaller than the LO phonon energy of GaAs. This value is also well outside any correction brought about by the warping of the hole bands. Such a puzzling phenomenon was also observed in the PLE spectra, from the bulk GaAs in which the oscillation began at photoexcitation energy less than the energy of one LO phonons.⁸ We believe the final state to which a hot electron can relax, is an exciton energy state rather than the bottom of the lowest conduction band. Obviously, this assumption was supported by the absolute absorption spectra of excitons.^{21,22} Under this assumption P_6 in Fig. 2(a) can be interpreted as follows: since the detection energy was set at the lower energy tail (1.563 eV) of the $1s\ n=1\ (e-hh)$ exciton state the energy difference between P_6 and the detection position is 43 meV, exactly the same excess energy loss of a hot electron by one LO phonon emission observed in bulk GaAs.⁸ This value is almost consistent with the calculation of $\hbar\omega(1+m_e/m_h)$ at the k value of $0.0165\ 2\pi/a$ as shown in Fig. 3. It seems that the P_6 most likely arises from the one LO phonon emission in the GaAs well, when the hot electrons photoexcited from the $n=1hh$ band at 1.606 eV relax to the detection energy position at 1.563 eV.

One plausible interpretation of P_7 in Fig. 2(a) is that the peak arises from one LO photon emission, when the hot electrons excited from the $n=1lh$ band at the k value of $0.0165\ 2\pi/a$ relax to the detection position of 1.563 eV. This means that hot electrons excited from both $n=1hh$ and lh valence band are at the same k value before relaxation. From Figs. 2(a) and 3, we can see that the energy difference between P_7 and P_6 is the same energy difference between $n=1lh$ and hh at $k=0.0165\ 2\pi/a$. It is noticed that the intensity of P_7 is weaker than that of P_6 . One reason of this phenomenon is that the optical transition probability of $n=1\ (e-hh)$ exciton states is greater than that of $n=1\ (e-lh)$ exciton states.²³ Further, when the detection energy was set at $2s\ n=1\ (el-hh)$ exciton line, the P_7 shifted to the position at ~ 1.635 eV as shown in Fig. 2(b). In Fig. 2(b) the position of P_7 relative to P_6 is also in the same range of the energy difference between $n=1lh$ and hh band at k value of $0.0231\ 2\pi/a$. An alternative interpretation of the P_7 in Fig. 2(a) may be from the result of one LO phonon emission of the hot electrons photoexcited from $n=1hh$ band, because P_7 just sits one LO phonon energy above the $1s\ n=1\ (e-hh)$ free exciton state. The PLE measurements show that when the detection energy was moved away from the peak of $1s\ n=1\ (el-hh)$ exciton line, the position and the intensity of P_7 changes. When the detection energy was

set at the $2s\ n=1\ (el-hh)$ exciton line which is 11 meV higher than the $1s$ exciton line the peak disappeared completely. Thus, we may have a conclusion that the formation of P_7 could be resulted from both of the electron relaxation processes.

In Fig. 2(a) a broad peak (P_8) around 1.625 eV is clearly resolved and in Fig. 2(b) the peak at around 1.625 eV overlaps P_6 as discussed above. However, when the detection energy of the PLE measurement was set at $1sn = 1\ (el-hh)$ exciton line (1.571 eV) the P_8 was unresolved. The formation of P_8 is possibly resulted from the combination of the following two transition processes: The first interpretation of the P_8 is due to the $n=1$ conduction band to the $n=2$ heavy-hole band transition. This assignment is in agreement with the $(1el-2hh)$ band gap calculation described above. However, the transition is an intrinsic effect and the PL intensity should be weak because it is a prohibition transition. Another possible explanation of the P_8 is the result of two-step relaxation processes: The first step is that the hot electrons photoexcited at 1.625 eV cascade down to the $2s\ n=1\ (e-hh)$ exciton states at 1.581 eV by one LO phonon emission. The subsequent processes are that some of the electrons at the $2s\ n=1\ (e-hh)$ states further relax to the monitored position at 1.563 eV by acoustic phonon emission and others recombine radiatively. Obviously, the carriers at the $2s\ n=1\ (e-hh)$ energy line have higher radiative recombination probability, than that of adjacent energy states as observed in both Figs. 1 and 2(a). Hence, when the detection energy is moved to the $2s\ n=1\ (e-hh)$ exciton line (1.581 eV) the position of the P_8 may not shift and the intensity of P_8 is relatively high as shown in Fig. 2(b).

The peak 10 at ~ 1.656 eV shown in Fig. 2(a) identifies with two subsequent LO phonon emissions, when the hot electrons excited from $n=1hh$ band cool down to emitting states: First, hot electrons at 1.656 eV relax to P_6 position at 1.606 eV. The energy loss of the hot electrons in this process is ~ 50 meV, ~ 7 meV greater than the energy loss of hot electrons in the bulk GaAs. The difference may arise from the nonparabolicity of the $n=1$ conduction band and the warping of heavy- and light-hole bands further away from the Γ point. The subsequent relaxation behavior of the hot electrons is in the same way as that of P_6 . This assignment was supported by the further PLE measurement shown in Fig. 2(b), in which the detection energy was moved to 1.581 eV, and the P_{10} shifted to the position at ~ 1.672 eV, two LO phonon energy higher than the $2s\ n=1\ (e-hh)$ exciton state.

Finally, we should note that in the above discussion we are only concerned with electron-phonon interaction. The reason for this is that for the purposes of this study we used the high purity sample. So carrier-ionized impurity interaction can be ignored. Also, in the PLE experiments low excitation density was used. Hence, carrier-carrier interaction was neglected as well.

In summary, with the use of photoluminescence excitation spectroscopy we have for the first time seen the oscillatory

behavior related to LO phonon emission in the GaAs/AlGaAs MQW structure. The period of the oscillation is different from that observed in bulk GaAs. The calculation from a four-band Kane model, describing the mixing of heavy- and light-hole bands at wave vector away from $k = 0$ was used to interpret the difference of oscillation features from the GaAs/AlGaAs MQW structure and bulk GaAs. We believe that the oscillation difference is due to the nonparabolicity and in-plane warping of the subbands in the MQW structure. The measured PLE spectra and the calculated results show that when the excess energy of a hot electron is equal to an integral number of the LO phonon energy of GaAs, the electron will directly cascade downwards to $1s$ $n=1$ ($e-hh$) exciton states by the LO phonon emission rather than to the bottom of $n=1$ conduction band. It is noted by the authors that there may be alternative interpretations of the peaks within higher energy region in Fig. 2, although the above interpretation is plausible.

ACKNOWLEDGMENTS

The authors are grateful to Z. Z. Li, Hangzhou University, for helpful discussions. This work has been supported in part by NSFC and the Office of State Education Commission.

¹U. Bockelmann and G. Bastard, Phys. Rev. B **42**, 8947 (1990).

²A. N. Forshaw and D. M. Whittaker, Phys. Rev. B **53**, 6955 (1996).

- ³H. Benisty, C. M. Sotomayor-Torres, and C. Weisbuch, Phys. Rev. B **44**, 10 945 (1991).
- ⁴C. H. Yang and S. A. Lyon, J. Cryst. Growth **134**, 305 (1985).
- ⁵E. O. Goebel and O. Hildebrand, Phys. Status Solidi B **88**, 645 (1978).
- ⁶P. D. Wang and C. M. Sotomayor-Torres, J. Appl. Phys. **74**, 5047 (1993).
- ⁷R. Notzel, L. Daweritz, N. N. Ledentsov, and K. Ploog, Appl. Phys. Lett. **60**, 1615 (1992).
- ⁸C. Weisbuch, Solid-State Electron. **21**, 179 (1978).
- ⁹R. E. Nahory, Phys. Rev. **178**, 1293 (1969).
- ¹⁰A. Chomette, B. Lambert, B. Clerjoud, F. Clerot, H. W. Liu, and A. Regreny, Semicond. Sci. Technol. **3**, 351 (1988).
- ¹¹R. C. Miller, D. A. Kleinman, W. T. Tsang, and A. C. Gossard, Phys. Rev. B **24**, 1134 (1981).
- ¹²M. H. Meynadier, C. Delalande, G. Bastard, M. Voos, F. Alexandre, and J. L. Lievin, Phys. Rev. B **31**, 5539 (1985).
- ¹³K. J. Moore, G. Duggan, P. Dawson, C. T. Foxon, N. J. Pulsford, and R. J. Nicholas, Phys. Rev. B **39**, 1219 (1989).
- ¹⁴K. J. Moore, P. Dawson, and C. T. Foxon, Phys. Rev. B **34**, 6022 (1986).
- ¹⁵R. White and L. J. Sham, Phys. Rev. Lett. **47**, 879 (1981).
- ¹⁶G. Bastard, Phys. Rev. B **24**, 5693 (1981).
- ¹⁷D. A. Broido and L. J. Sham, Phys. Rev. B **34**, 3917 (1986).
- ¹⁸Y. C. Chang and J. N. Schulman, Phys. Rev. B **31**, 2069 (1985).
- ¹⁹M. F. H. Schuurmans and G. W. 'tHooft, Phys. Rev. B **31**, 8041 (1985).
- ²⁰R. Eppenga, M. F. H. Schuurmans, and S. Colak, Phys. Rev. B **36**, 1554 (1987).
- ²¹R. Dingle and W. Wiegmann, Phys. Rev. Lett. **34**, 1327 (1975).
- ²²J. E. Zucker, P. S. Chemla, A. Gossard, and W. Wiegmann, Second International Conference on Modulated Semiconductor Structures, 1985 (unpublished), p. 96.
- ²³R. C. Miller, A. Gossard, G. D. Sanders, and Y. C. Chang, Phys. Rev. B **32**, 8452 (1985).

Base metallization stability in InP/InGaAs heterojunction bipolar transistors and its influence on leakage currents

D. Caffin,^{a)} C. Besombes, J. F. Bresse, P. Legay, G. Le Roux, G. Patriarche, and P. Launay
France Telecom, CNET-PAB, Laboratoire de Bagneux, BP 107, 92225 Bagneux Cedex, France

(Received 25 June 1996; accepted 18 April 1997)

For InP/InGaAs heterojunction bipolar transistors (HBTs), base-collector leakage current can be quite impairing by restricting their operational conditions to a very narrow emitter-collector voltage range. Results of a study on morphological degradations of various *p*-type metallizations to InGaAs and their effect on base-collector leakage current of InP/InGaAs double HBTs are reported here. Two kinds of base contacts were investigated. Mn/Au/Ti/Au induces high leakage current after contact annealing at temperatures as low as 300 °C. This is due to important interdiffusion of the species, and precipitation inside the base layer. Ti/Pt/Au is a good alternative, provided that the platinum layer is not too thick, and the contact annealing temperature not too high. Leakage current was found to be related to the strain induced by metal deposition, or to the morphological degradation taking place during annealings at high temperatures (metal precipitates formation inside the base layer inducing strain, and III-V elements exodiffusion). However, HBTs with Ti/Pt/Au base contacts annealed at 300 °C showed no excess leakage current and almost no morphological alteration, and contact resistances below $10^{-5} \Omega \text{ cm}^2$ were obtained for base doping levels above 10^{19} cm^{-3} . © 1997 American Vacuum Society. [S0734-211X(97)06004-6]

I. INTRODUCTION

In the race for very high-speed microelectronic and optoelectronic devices, InP/InGaAs heterojunction bipolar transistors (HBTs) have emerged as a promising bipolar technology. Very high performance microwave InP/InGaAs HBTs have already been reported.¹ For such devices, the base metallization is an important issue, because the base contact resistance can be a very limiting factor. Various *p*-type metallizations have been used, falling into two main categories: contacts including a diffusing *p*-type dopant (such as Zn, Be, or Mn), or nonalloyed contacts (such as Ti/Au, Ti/Pt/Au, or Pt/Ti/Pt/Au). The former, commonly used for GaAs based HBTs, provide a highly doped interfacial layer, but lead to important inhomogeneities and poor thermal stability;^{2,3} the latter have become the most popular contacts for InP based HBTs over the past few years. Very low contact resistances have been reported for such metallizations.^{3,4}

An important problem encountered with InP/InGaAs HBTs is the reverse leakage current of the base-collector junction, which may become large well before junction breakdown. This effect limits the practical use of the device to a very narrow emitter-collector voltage region. The influence of the base metallization on leakage currents is investigated in article: upon contact deposition or annealing, morphological degradation of the base layer beneath the metal may appear, and even reach the base-collector interface, thus, considerably degrading the junction characteristics. Very little work on this problem has been reported so far. However, this effect can be all the more critical as the base becomes thinner, and can even lead to short circuits, if the metallization is ill chosen.⁵ Section III presents the influence

of a Mn/Au/Ti/Au contact, which was successfully used for GaAs HBTs.⁶ In Sec. IV, various Ti/Pt/Au structures are investigated.

II. EXPERIMENTS

The HBT structure used in this study to evaluate the leakage currents is described in Table I. The samples were grown by chemical beam epitaxy (CBE) on Fe-doped semi-insulating substrates. The dopant sources were solid silicon (*n*-type), and CBr_4 (*p*-type).

Large devices were realized as follows: The emitter was etched chemically; the base mesa was defined by a combination of wet etching and ion milling in order to prevent base-collector leakage related to the etching process;⁷ device insulation was performed by wet etching. Metallizations were realized by conventional lift-off techniques. For wet etchings, the etching mixtures are $\text{H}_3\text{PO}_4/\text{H}_2\text{O}_2/\text{H}_2\text{O}$ (3/1/40) for InGaAs and $\text{HCl}/\text{H}_3\text{PO}_4$ (1/10) for InP.

Ti/Pt/Au was used for low resistance *n*-type Ohmic contacts (around $2 \times 10^{-7} \Omega \text{ cm}^2$). Mn/Au/Ti/Au and Ti/Pt/Au (with various platinum thicknesses) were tested to form the *p*-type base Ohmic contact. All metals were deposited by electron-beam evaporation (with a background pressure below 10^{-7} Torr), and annealed in a rapid thermal annealer, under Ar/H_2 flow. The temperature ramp was 30 s long for all annealings, and the samples were rapidly cooled down by a nitrogen "shower." The emitter-base and base-collector areas were 170×100 and $235 \times 190 \mu\text{m}^2$, respectively.

All *I*-*V* characteristics were measured with a HP4145 A parameter analyzer on at least five transistors. In order to evaluate the excess leakage currents created by metal depo-

^{a)}Electronic mail: david.caffin@bagneux.cn

TABLE I. HBT general structure.

Layer	Material	Thickness (Å)	Doping (cm ⁻³)
Emitter contact 1	<i>n</i> ⁺ InGaAs	1000	7.10 ¹⁸ –2.10 ¹⁹
Emitter contact 2	<i>n</i> ⁺ InP	500	1–3.10 ¹⁹
Emitter	<i>N</i> InP	1500	10 ¹⁷ –5.10 ¹⁷
Base	<i>p</i> ⁺ InGaAs	1000	8.10 ¹⁸ –2.10 ¹⁹
Collector	InGaAs	200	Undoped
	InGaAsP (<i>E_g</i> = 0.95 eV)	200	Undoped
	InGaAsP (<i>E_g</i> = 1.15 eV)	200	Undoped
	<i>n</i> ⁻ InP	5000	10 ¹⁶
Collector contact 1	<i>n</i> ⁺ InGaAs	2000–5000	7.10 ¹⁸ –2.10 ¹⁹
Collector contact 2	<i>n</i> ⁺ InP	3000–0	1–3.10 ¹⁹

sition or contact annealing, some devices were left without any base contact. Unfortunately, it is then impossible to measure the reverse current of the base–collector diode. Instead, the collector current versus emitter–collector voltage characteristic (for $I_b = 0$) was measured in order to evaluate the leakage current of the whole HBT structure.

III. Mn/Au/Ti/Au

Mn/Au/Ti/Au provides a low resistance Ohmic contact by diffusion of the manganese (which is a *p*-type dopant) into the base layer. The thicknesses of the layers were 300/2000/300/1000 Å for Mn/Au/Ti/Au. The sample was annealed at various temperatures for 5 min. Figure 1 shows the evolution of the emitter–collector leakage current at $V_{ce} = 4$ V with the annealing temperature. Above 250 °C, the leakage current is much higher for the device with the base contact than for the

one without, showing that the leaks are generated by the interactions of the metallization with the semiconductor layers upon annealing.

Figure 2 shows secondary ion mass spectroscopy (SIMS) profiles of Au/Ti/Au/Mn/InGaAs/InP samples. The epitaxial structure, grown by metal–organic chemical–vapor deposition, simulates the base–collector junction (the InGaAs layer is 1400 Å thick). One reference sample was left unannealed; the second and third ones were annealed at 300 °C for 5 min, and 450 °C for 25 s, respectively. The last one was realized at a rather high temperature in order to enhance the phenomenon taking place upon annealing. Auger analysis were also performed in order to have a quantitative glimpse at the relative concentrations of the species at stake. The Auger profiles of the samples annealed at 300 °C and 450 °C are presented in Figs. 3(a) and 3(b), respectively. They were corrected with the elemental sensitivity factors. Thus, the Auger relative intensities give a good estimation of the relative atomic concentrations (with a relative accuracy better than 20%).

Large intermixing of the various elements occurs upon annealing. Indium diffuses upward through the metal layers, but is partially blocked by the titanium as can be seen on the SIMS profiles. This diffusion from the semiconductor is important even after the low-temperature annealing (300 °C): the corresponding Auger profile shows that the InGaAs layer is already In poor, although the indium concentration inside the metal layers is too low to appear on this profile. However, it starts showing on the 450 °C profile.

Arsenic also diffuses upward, and accumulates into the Ti layer. After the 300 °C anneal, the diffusion is not very important, and only shows on the SIMS profile. After the 450 °C annealing, the Auger profiles reveal that the arsenic/titanium mixture is almost stoichiometric. The exodiffusion of indium and arsenic has converted the InGaAs layer into a material very close to GaAs.

Some phosphorous diffuses towards the surface through the whole structure, with a slight accumulation into the InGaAs layer. At 450 °C, more diffusion takes place, and phosphorous is also partially blocked by the Ti layer.

As for the metal diffusion, the SIMS profiles are harder to analyze: the metal signal tails in the layers below are partly a SIMS artefact, due to etching nonhomogeneity. However,

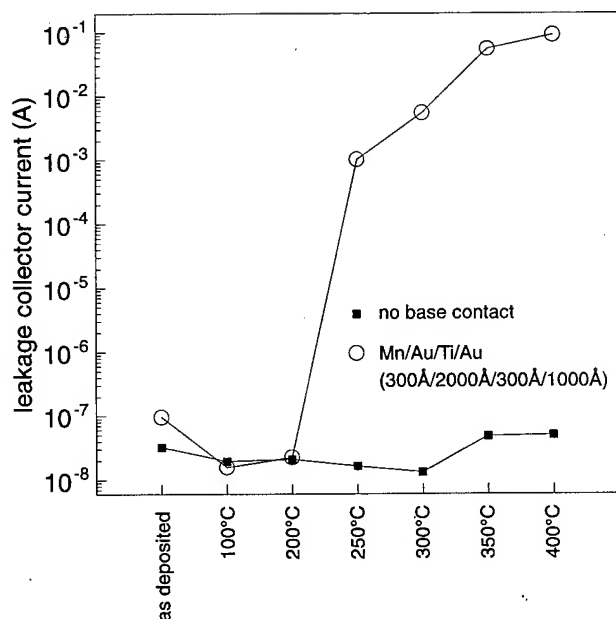


FIG. 1. Transistor leakage currents at $V_{ce} = 4$ V vs annealing temperature for the Mn/Au/Ti/Au base contact.

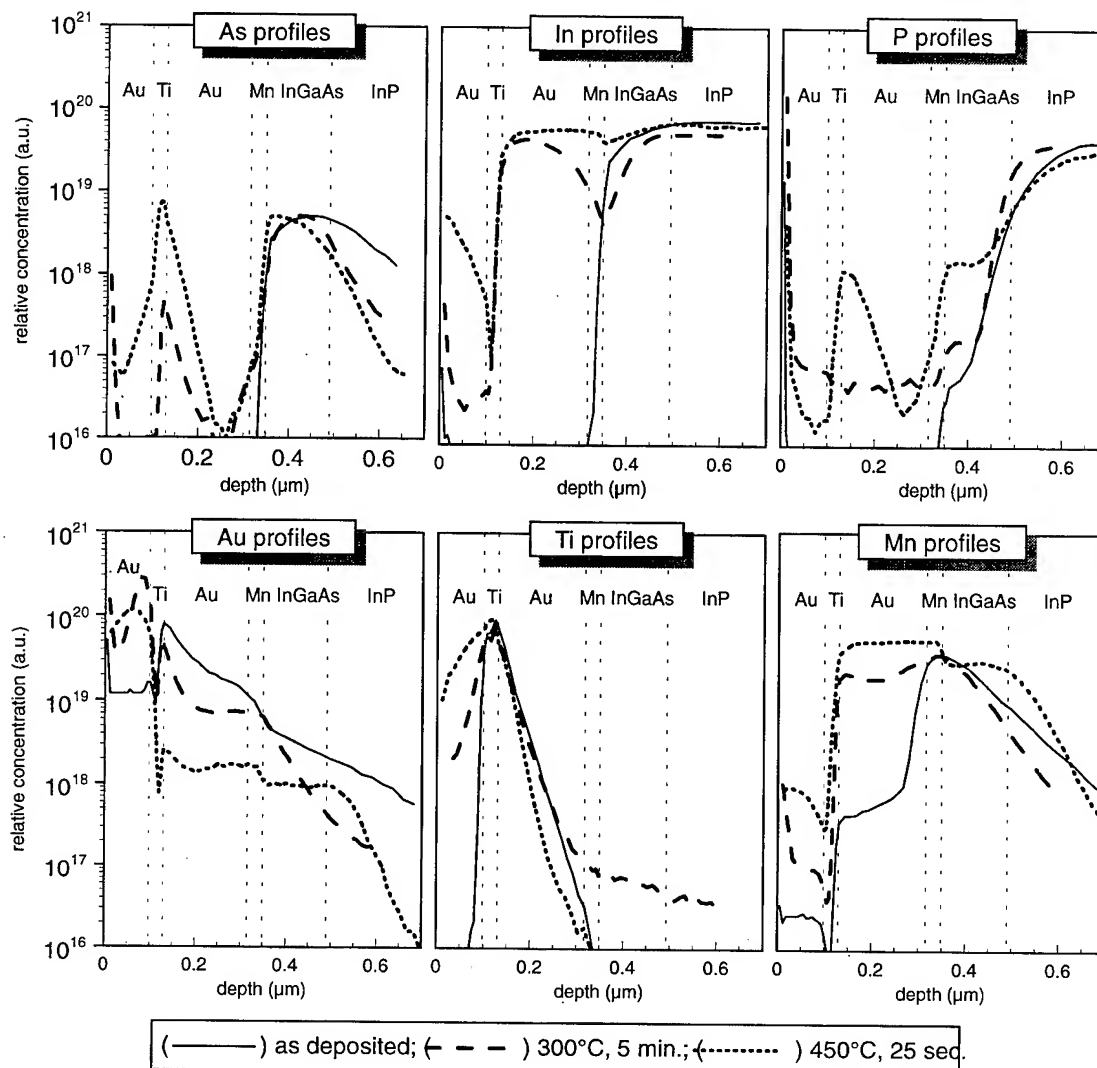


FIG. 2. SIMS profiles of the Au/Ti/Au/Mn/InGaAs/InP structure annealed at various temperatures.

information can be obtained by comparing one set of curves to another. Titanium is rather stable except for some diffusion into the top gold layer: with the penetration of arsenic inside the layer, the titanium atoms are "pushed" into the gold layer above.

The evolution of manganese and gold is given more clearly by the Auger profiles than by the SIMS data. However, it is worth noticing that the Mn SIMS profile of the as-deposited sample reveals the presence of manganese inside the gold layer above, even before annealing, most probably due to diffusion during the gold and titanium deposition. After the 300 °C anneal, the Auger profile shows that the Mn layer has almost disappeared. Some of the manganese has diffused into the semiconductor below (as expected and desired for such a contact), but most of it has diffused upward. It seems that in the process, there is an exchange between Au and Mn atoms, with gold diffusion into the InGaAs below.

Figure 4 presents a cross-sectional transmission electron microscopy (TEM) micrograph of the sample annealed at

300 °C. Large precipitates in the InGaAs layer can be seen, showing that the gold atoms that have diffused from the layer above, have precipitated inside the semiconductor. However, it is worth noticing that scanning electron microscope observations of the metallization surface after annealing did not show morphological degradation.

After the 450 °C annealing, the manganese layer has totally disappeared. The gold layer above is Mn rich, as shown by the plateau of the Mn profile in Fig. 3(b), and the base layer is saturated with gold. Manganese has also diffused far into the semiconductor layers below (more than 1000 Å into the InP collector).

X-ray diffraction spectra were realized to check whether new phases were formed (Fig. 5). After the 450 °C annealing, two diffraction lines appear around the Au 111 reflection [Fig. 5(a)]. According to the American Society for Testing and Materials (ASTM) tables, these lines could be attributed to an alloy close to TiAs. Figure 5(b) presents the diffraction spectra for an angle range corresponding to the tail of the 004 InP substrate peak. After the 300 °C annealing, the tail

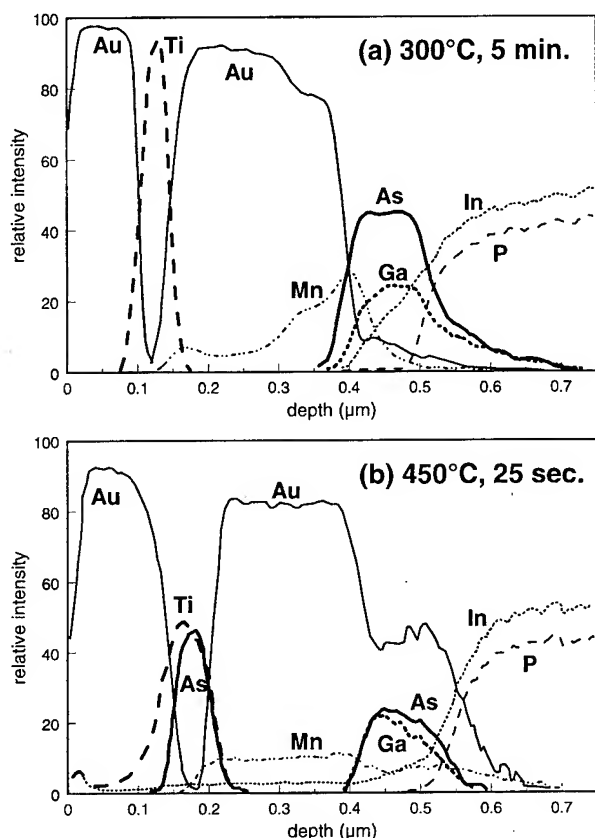


FIG. 3. Auger depth profiles of the Au/Ti/Au/Mn/InGaAs/InP structure annealed at 450 °C for 25 s.

broadens, and after the 450 °C annealing, a new diffraction line appears around $2\theta=66^\circ$. This value is close to the 004 reflection of GaAs. This corresponds well to the indium out-diffusion into the metal, leading to indium impoverishment of the InGaAs layer.

From all these results, it appears that upon annealing, the base layer and base-collector interface are very much degraded by the interdiffusions of the species at stake, and it is no wonder the diode leakage current becomes very important. Early work on $p-n$ silicon diodes reported that crystalline defects or precipitates can induce leakage currents, arising from local enhancement of the electric field, and local Zener effect.^{8,9}

IV. Ti/Pt/Au

Ti/Pt/Au is commonly used to realize a low resistance nonalloyed p -type Ohmic contact. The titanium layer provides a good adhesion of the contact onto the semiconductor, and the platinum layer acts as a diffusion barrier to ensure good stability. Katz and co-workers widely studied Pt/Ti contacts on p -InGaAs,^{4,10} reporting lowest contact resistance (in the $10^{-8} \Omega \text{ cm}^2$ range⁴) after a rapid thermal annealing at 450 °C. For this metallization, they found that the morphology of the contact was not modified for annealing temperatures up to 400 °C, and that a slight intermixing of the semiconductor and metal layers took place at 450 °C. Stareev *et al.* even found that the structure InGaAs/Ti/Pt/Au was stable up to 500 °C.³ However, we find that, for HBTs, Ti/Pt/Au can induce important base-collector leakage currents in this range of temperature.

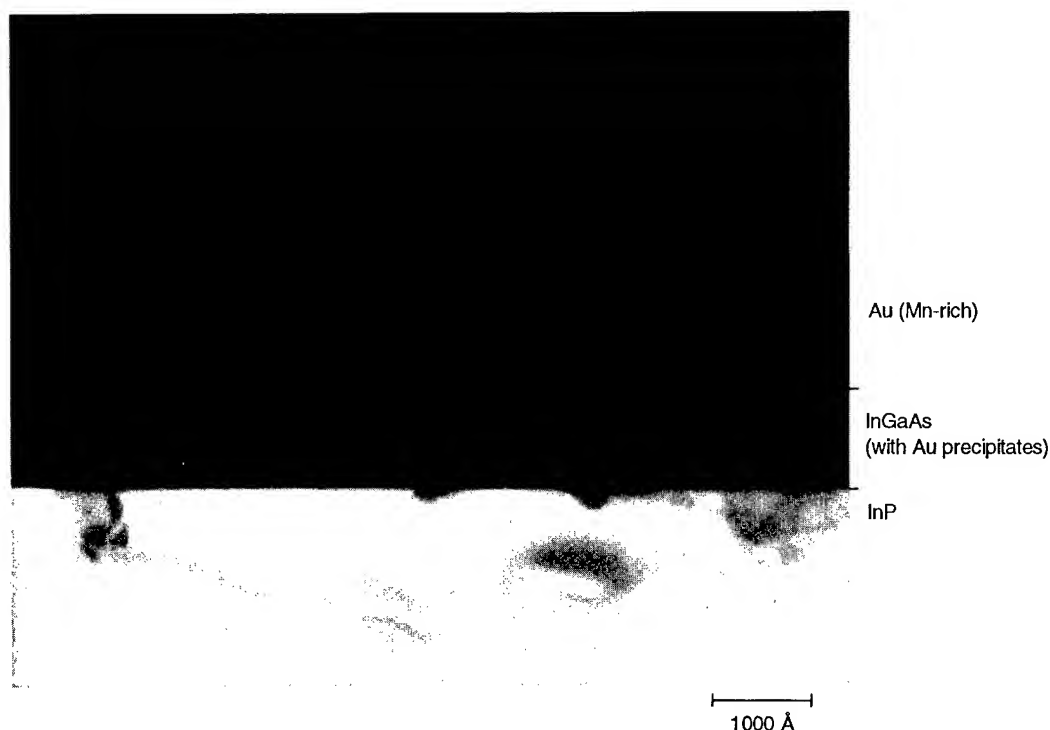


FIG. 4. TEM micrograph of the Au/Ti/Au/Mn/InGaAs/InP structure annealed at 300 °C for 5 min.

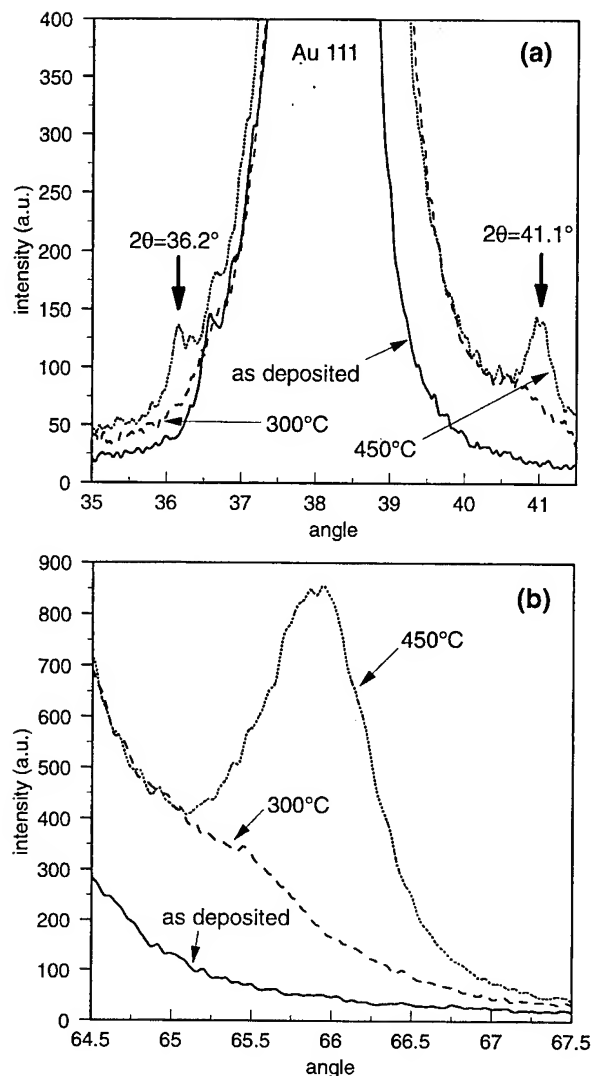


FIG. 5. X-ray diffraction spectra of the Au/Ti/Au/Mn/InGaAs/InP structure annealed at various temperatures.

A. Influence of metal deposition

In Fig. 6, the transistor leakage current at $V_{ce} = 4$ V is reported for various Ti/Pt/Au contact thicknesses. Leakage currents with no base contact are also reported. Excess leakage is created by the metal deposition (before annealing), increasing with the platinum thickness. This is most probably due to the strain generated during platinum deposition. Strain measurements were performed by an optical method. With this method, the strain is obtained by measuring the curvature of the sample before and after metal deposition. It is rather awkward to obtain reliable results with epitaxial layers, so the values reported here were measured on InP substrates. The strain measured for Ti/Pt/Au, 300/500/2500 Å, was 145 MPa, more than three times higher than that obtained with Ti/Au, 500/2500 Å (42 MPa). Since as-deposited measurements reported in Fig. 6 were performed with large unannealed contacts, it can be concluded that it is only the strain in the base layer that generates the leaks. This was

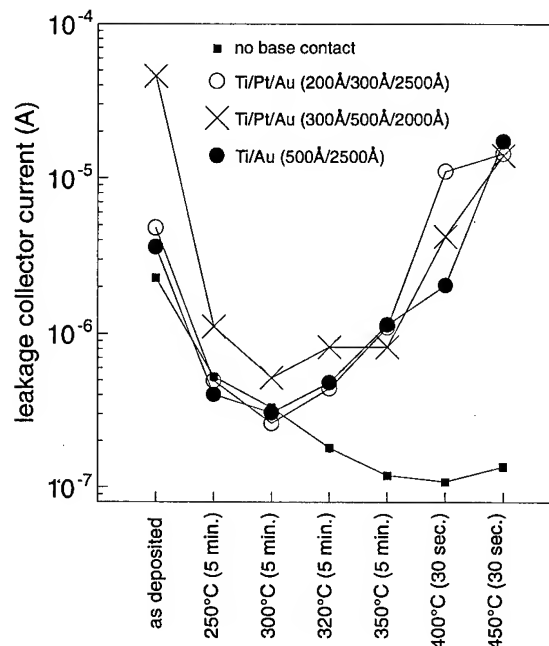


FIG. 6. Transistor leakage currents at $V_{ce} = 4$ V vs annealing temperature for various base contacts.

confirmed by x-ray diffraction measurements, as will be discussed in the following. It is worth noticing that small devices should be less sensitive to the strain induced by the metal.

B. Influence of contact annealing

Figure 6 shows that, for annealing temperatures below 300 °C, leakage decreases with increasing temperature. This is probably due to strain relaxation, and epitaxial defects curing (as this leakage reduction also appears for the sample with no base contact). For heat treatments at a temperature higher than 300 °C, leakage current drastically increases for the samples with base contacts, when it goes on decreasing slowly for the sample with no base contact. This demonstrates that leakage is generated by the interactions of the metal and semiconductor layers upon annealing. The higher the temperature, the larger the degradation.

SIMS analysis were performed on Au/Pt/Ti/InGaAs/InP samples. The thicknesses were Au: 2000 Å, Pt: 500 Å, Ti: 300 Å, and InGaAs: 1400 Å. The profiles are presented in Fig. 7 for three annealing conditions: not annealed, 300 °C (5 min), and 450 °C (25 s).

After the 300 °C treatment, some indium has diffused from the base, and accumulates into the platinum layer. At 450 °C, indium goes through the platinum, into the gold layer. Arsenic roughly follows the same pattern as indium, but much attenuated: the 450 °C As profile looks like the 300 °C In profile. Some phosphorous also diffuses into the layers above, and at 450 °C, goes through the whole metallization to reach the surface of the sample. As for the metals, SIMS profiles give three pieces of information: upon anneal-

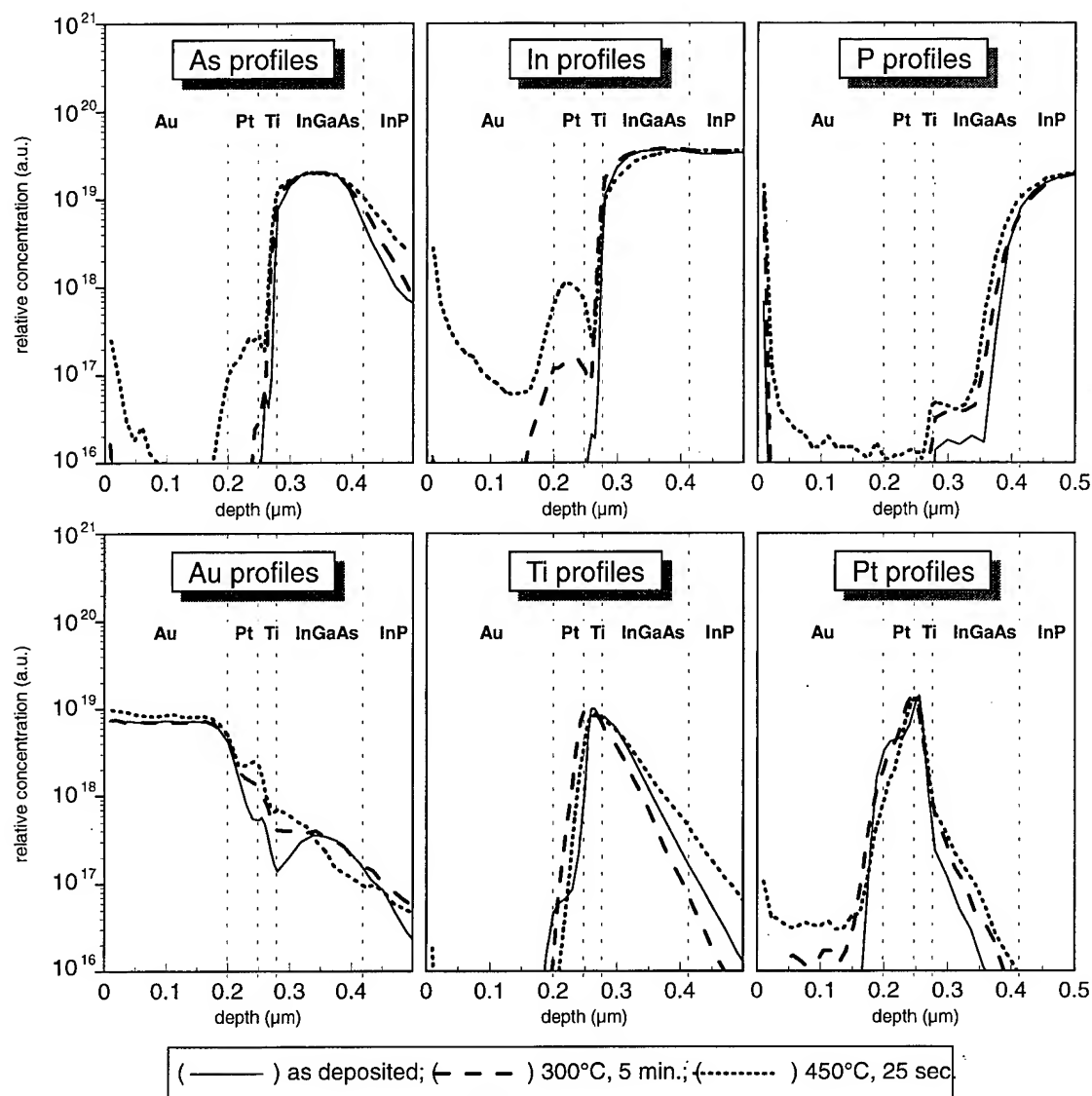


FIG. 7. SIMS profiles of the Au/Pt/Ti/InGaAs/InP structure annealed at various temperatures.

ing, the gold layer becomes platinum rich, some gold diffuses into the layers below, and the titanium seems rather stable.

X-ray diffraction analyses were realized to identify possible new phases. No extra line was found after annealing, but displacements of the metal peaks were observed. Gold and platinum both have a cubic, face-centered structure, and the 111 lines are the more intense lines obtained on our spectra, showing that the two metals are deposited with a (111) preferential orientation. Figure 8 presents the 111 diffraction lines for gold [Fig. 8(a)] and platinum [Fig. 8(b)] before and after heat treatments. The corresponding lattice parameters are indicated. The unannealed gold layer lattice constant corresponds to that given by the ASTM tables ($a = 4.079 \text{ \AA}$), but for platinum, the constant is smaller than that of the ASTM tables ($a = 3.923 \text{ \AA}$), which reveals a strained layer. For both metals, the 111 peak intensity is much higher after annealing, indicating the regrowth of

(111)-oriented grains. After annealing, the platinum lattice parameter is almost the one given by the ASTM tables, indicating that the layer has relaxed. On the other hand, the gold lattice constant decreases upon annealing. This is rather difficult to understand, but might be due to the penetration of other elements (such as indium or platinum) into the gold layer, thus, compressing the gold lattice.

Cross-sectional transmission electron microscopy of the films was also performed in order to have a better understanding of the mechanisms at stake. In Fig. 9, a TEM micrograph of the sample annealed at 450°C is shown (the 300°C sample presented no sensible morphological degradation). The intermixing of the gold and platinum layers observed by SIMS is actually identified as a growth of platinum grains into the gold layer, and vice versa. This could also account for the gold lattice parameter reduction (but not entirely, because such a reduction was also observed with Ti/Au contacts). Gold has precipitated into the top half of the

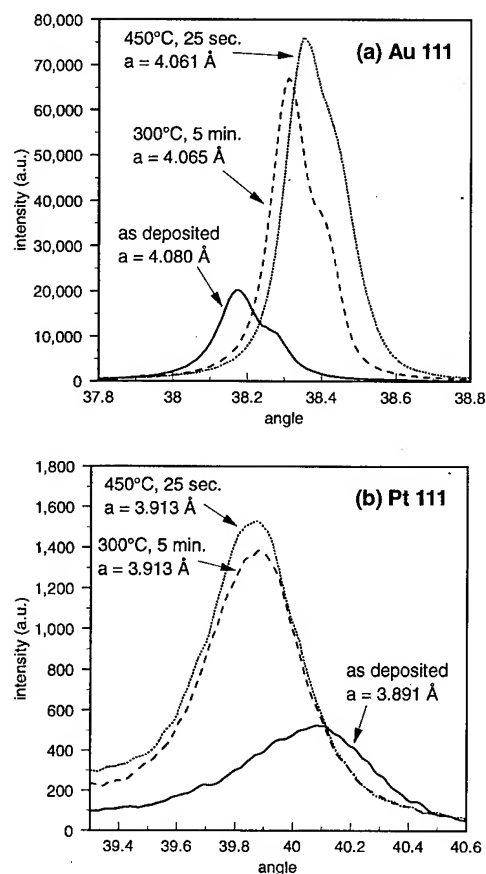


FIG. 8. X-ray diffraction spectra of the Au/Pt/Ti/InGaAs/InP structure annealed at various temperatures.

InGaAs layer. The precipitates have a tabular shape, and can be as thick as 600 Å. Smaller, scarcer, and rounder precipitates are also observed, and identified as PtAs_2 by x-ray element analysis. The fringes in the remaining base layer are due to the strain generated by the precipitates. The diffusion of gold into the base layer is compensated by indium and arsenic outdiffusion into the metal layers as shown in the SIMS profiles. These interdiffusions can even leave voids at the semiconductor/metal interface (white triangle in Fig. 9).

All these results are in good agreement with the leakage current measurements. No sensible morphological changes of the base layer appears after a 300 °C heat treatment. On the contrary, it allows relaxing the strain created by the metal deposition, and eventually curing some defects of the layers. Thus, the leakage current is lower after the 300 °C annealing than before. However, a high annealing temperature (450 °C) generates a huge amount of defects, precipitates, and strain in the base layer.

In order to achieve a better thermal stability of the metallization upon annealing, several alternate structures have been proposed to contact InGaAs. The refractory silicide WSi has been shown to be rather stable at annealing temperatures up to 800 °C, when Ti/Pt/Au failed at 400 °C.¹¹ A thin layer of tungsten inserted between the InGaAs layer and the Ti/Pt/Au structure was also shown to stabilize the metallization upon annealing at 400 °C, thus, reducing the leakage current of InP-based HBTs.⁵ Another refractory metal, Mo, could also be a good candidate for realizing thermally stable contacts to InGaAs. Mo/InGaAs structures have been found to be more stable than various other metal/semiconductor structures (such as Pt/InGaAs or Ti/InGaAs).¹² Furthermore, a comparative study on the thermal stability of Mo/Au and

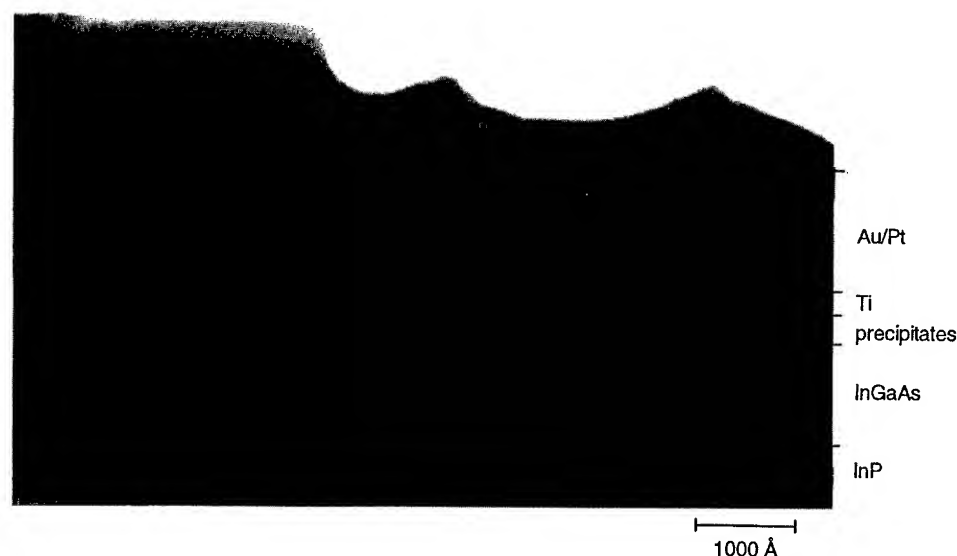


FIG. 9. TEM micrograph of the Au/Pt/Ti/InGaAs/InP structure annealed at 450 °C for 25 s.

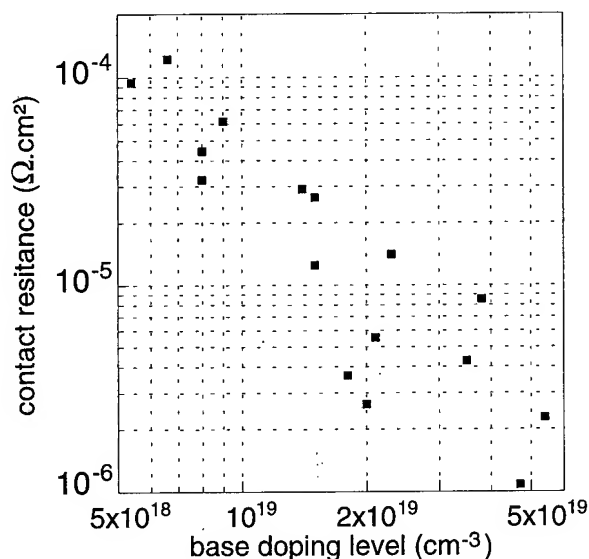


FIG. 10. Contact resistance vs base doping level for the Ti/Pt/Au contact (200/300/2500 Å) annealed at 300 °C for 10 min.

Ti/Pt/Au on *n*-type InGaAs showed that the former was more reliable over time at an annealing temperature of 300 °C.¹³

C. Contact resistance

All the 300 °C results presented till now are for 5 min annealings, but we found that the leakage current is stable when the samples are annealed for five more minutes. The contact resistance of the Ti/Pt/Au (200/300/2500 Å) metallization was measured by the transmission line method after 300 °C annealings. Lower contact resistances were measured after 10 min long annealings, but stabilized for longer annealings. Figure 10 presents the contact resistance for various base doping levels. Reasonably low values can be obtained, provided that the base is doped to a high enough level.

V. CONCLUSION

The morphological degradation of various metallizations on *p*-type InGaAs and its effect on base-collector leakage current of InP/InGaAs heterojunction bipolar transistors were investigated. The results reported here were obtained for a base-collector heterojunction, but a homojunction would most probably lead to the same kind of degradation.

The Mn/Au/Ti/Au Ohmic contact was found to create important leakage after heat treatment at temperatures as low as 300 °C. X-ray diffraction spectra, TEM observations, SIMS,

and Auger depth profiles revealed that large interdiffusion of the various elements takes place upon annealing. Arsenic and indium diffuse upward into the metal; the whole manganese layer is redistributed into the layers above and below; and gold precipitates into the base. New phases are even formed after a 450 °C annealing. Mn/Au/Ti/Au is, thus, unfit for contacting the base of HBTs.

A more stable, not alloyed *p*-type contact such as Ti/Pt/Au should be preferred. The platinum layer must not be too thick to avoid strain in the base layer. After a 300 °C, 10 min heat treatment, no leakage is induced by the metallization, but if the annealing temperature is higher, morphological degradation of the layers occurs, and the base-collector diode starts leading. X-ray diffraction, SIMS profiles, and cross-sectional TEM showed that 300 °C heat treatments preserve the layer integrity, but that higher temperatures induce degradation: gold precipitates inside the base layer, which, in turn, becomes strained and depleted in indium and arsenic.

Contact resistances in the 10^{-6} Ω cm² range were obtained with Ti/Pt/Au (200/300/2500 Å) annealed at 300 °C for 10 min for a base doping level larger than 2×10^{19} cm⁻³. Finally, Ti/Pt/Au is a good candidate for HBT's base metallization, provided that subsequent annealing temperatures stay below 300 °C.

ACKNOWLEDGMENTS

The authors would like to thank J. L. Benchimol and F. Alexandre for CBE growths, and M. Juhel for his help and support for the SIMS analysis.

- ¹Y. Matsuoaka and E. Sano, *Solid-State Electron.* **38**, 1703 (1995).
- ²J.-F. Thiery, H. Fawaz, A. Leroy, and G. Salmer, *J. Vac. Sci. Technol. B* **13**, 2130 (1995).
- ³G. Stareev, H. Künzel, and G. Dortmann, *J. Appl. Phys.* **74**, 7344 (1993).
- ⁴A. Katz, W. C. Dautremont-Smith, S. N. G. Chu, P. M. Thomas, L. A. Koszi, J. W. Lee, V. G. Riggs, R. L. Brown, S. G. Napholtz, J. L. Zilko, and A. Lahav, *Appl. Phys. Lett.* **54**, (1989).
- ⁵E. F. Chor, R. J. Malik, R. A. Hamm, and R. Ryan, *IEEE Electron Device Lett.* **17**, (1996).
- ⁶P. Launay, R. Driad, J. L. Benchimol, F. Alexandre, and L. Dangla, *ESSDERC 94 Proceedings*, pp. 439–442, 1994.
- ⁷D. Caffin, A. M. Duchenois, F. Hélot, C. Besombes, and P. Launay, *IEEE Trans. Electron Devices* **44**, (1997).
- ⁸A. Goetzberger and W. Shockley, *J. Appl. Phys.* **31**, 1821 (1960).
- ⁹W. Shockley, *Solid-State Electron.* **2**, 35 (1961).
- ¹⁰A. Katz, S. N. G. Chu, B. E. Weir, C. R. Abernathy, W. S. Hobson, S. J. Pearton, and W. Savin, *IEEE Trans. Electron Devices* **39**, (1992).
- ¹¹T. Nittono, H. Ito, O. Nakajima, and T. Ishibashi, *Jpn. J. Appl. Phys.* **1** **27**, 1718 (1988).
- ¹²Y. Ashizawa, C. Nozaki, T. Noda, and A. Sasaki, *IPRM* **95**, 861 (1995).
- ¹³A. S. Wakita, N. Moll, S. J. Rosner, and A. Fisher-Colbrie, *J. Vac. Sci. Technol. B* **13**, 2092 (1995).

Structural and optical characterizations of single three-dimensionally confined GaAs/AlAs structures grown on patterned GaAs (001) substrates

Jun-ichi Kasai,^{a)} Sige-hisa Tanaka, and Katsuhiko Higuchi
Central Research Laboratory, Hitachi, Ltd., Kokubunji, Tokyo 185, Japan

Yoshifumi Katayama^{b)}
Optoelectronics Technology Research Laboratory, 5-5 Tohkodai, Tsukuba, Ibaraki 300-26, Japan

(Received 16 January 1997; accepted 16 May 1997)

We report structural and optical characteristics of single three-dimensionally (3D) confined GaAs/AlAs structures grown by molecular beam epitaxy on $\langle 100 \rangle$ oriented square mesas patterned onto GaAs (001) substrates. By fabricating single structures and using AlAs barrier layers, we were able to clearly attribute luminescent spectral peaks to spatial origins in the grown structures. Observation with a scanning electron microscope and an atomic force microscope indicated that the structures on the mesas were pinched-off or truncated pyramids limited by sidewall $\{011\}$ facets and a mesa-top facet. The observation also led to estimation of lateral width in the 3D confined structures, which decreased in proportion to a decrease in pattern width ranging from 1.6 to 0.2 μm . This proportionality can be explained by adatom migration on the sidewalls towards the bottom region of the patterned substrates. The optical properties of the grown structures were investigated by microphotoluminescence measurements at 8 K. Photoluminescence (PL) images were first measured to attribute luminescent spectral peaks to quantum-well (QW) heterostructures spatially originating in a truncated pyramidal structure. We next performed spectroscopy on single 3D confined structures of various lateral widths by positioning and holding the mesa top of one isolated pyramidal structure at a time under laser illumination. The PL and PL excitation spectra exhibited distinct exciton peaks about 10 meV wide. As the lateral width decreased, the vertical layer thickness estimated by their emission wavelength increased, but it saturated at smaller lateral widths. The saturation can also be explained by the adatom migration on the sidewalls. When excitation power was lowered, PL spectra became a cluster of sharp peaks. Each sharp peak may be due to the recombination of the exciton trapped in QW layer thickness fluctuation. The cluster thus suggests that the broad spectral width is caused mainly by wide variation in QW layer thickness of 3D confined structures. © 1997 American Vacuum Society. [S0734-211X(97)13604-6]

I. INTRODUCTION

Fabrication of nanostructures is one of the main subjects in recent semiconductor research.^{1,2} Controlling band gap patterns in two or three dimensions allows spatial confinement of charge carriers to small volumes. For extremely small confining potential wells, on the nanometer scale, two- or three-dimensional quantum effects become significant. These extremely small confined structures, i.e., nanostructures, are expected to exhibit new electrical and optical properties that can be used for novel device applications.

Various attempts have been made to fabricate semiconductor nanostructures.¹⁻⁴ One fabrication technique is based on selective growth on patterned substrates by using epitaxial growth methods such as molecular beam epitaxy (MBE) or metalorganic chemical vapor deposition (MOCVD).⁴ This technique uses the difference in growth rate on neighboring facets and can grow two-dimensionally (2D) or three-dimensionally (3D) confined structures during the epilayer growth process itself. All interfaces of the confining potential

wells are formed *in situ* and can be free of defects and contamination.

3D confined structures are formed by growth on mesa structures whose top facet is surrounded by crystallographically equivalent side facets. For example, in GaAs (001) substrates, which are most often used for device applications, the top facet of $\langle 100 \rangle$ oriented square mesas is surrounded by equivalent sidewalls of $\{011\}$ facets. Group-III atoms arriving at the sidewall facets migrate towards the adjacent facet of the mesa-top plane. This adatom migration results in a smaller growth rate on the sidewalls than on the mesa top and in shrinkage of the mesa-top area, which leads to mesa pinch off. The pinched-off mesa structures on patterned GaAs (001) substrates are therefore pyramidal ones surrounded by sidewall $\{011\}$ facets.

Growth of quantum-well (QW) heterostructures on patterned substrates brings about lateral variations in layer thickness and alloy composition. 3D confined structures are thus formed by growing QW heterostructures on pyramidal structures before the pinch off. The vertical confinement energy in the 3D confined structures is due to quantum confinement introduced by the QW heterostructures normal to their interface. The strong lateral-position dependence of this vertical confinement energy leads to lateral variation in va-

^{a)}Electronic mail: Junichi-cr1.Kasai@c-net3.crl.hitachi.co.jp

^{b)}Present address: Femtosecond Technology Research Association, 5-5 Tohkodai, Tsukuba, Ibaraki 300-26, Japan.

lence and conduction band edges. The lateral thickness variations thus result in effective band gap modulation. The thicker QW sections constitute lateral potential wells that can trap charge carriers, while the thinner QW sections function as effective lateral barriers, analogous to the higher band gap layers that serve as the vertical barriers in conventional QW heterostructures.

There have been several reports of 3D confined structures fabricated on patterned substrates.⁵⁻¹¹ Madhukar *et al.* obtained evidence for 3D confined structures with lateral sizes down to about 50 nm by observing cross sections of pyramidal structures with a transmission electron microscope (TEM).^{5,8} Cathodoluminescence (CL) images and the spectra of the 3D confined structures indicated their high optical quality.^{6,11} However, they grew GaAs/AlGaAs multiple QWs on patterned GaAs substrates in order to follow the evolution of the growth front. Although the growth time for the QW and barrier layers was kept constant, a strong increase in layer thickness on the mesa top occurred as growth proceeded, i.e., when the mesa-top width of the pyramidal structures decreased. Thus, owing to difference in the QW layer thickness, CL spectra exhibited some peaks due to the emissions from the various QW layers on the mesa top. This multiplicity made it difficult to attribute the luminescent spectral peaks to QW layers spatially originating in the pyramidal structure. Moreover, AlGaAs barrier layers in the QW heterostructure further complicated the attribution because of the lateral variation in alloy composition.^{12,13}

In order to simplify the attribution of luminescent spectral peaks to spatial origins, we fabricated single 3D confined structures having AlAs barrier layers instead of multiple ones having AlGaAs barrier layers. In this article we report structural and optical characteristics of the single 3D confined structures grown by MBE on patterned GaAs (001) substrates. Observation of the surface of the pyramidal structures with a scanning electron microscope (SEM) and an atomic force microscope (AFM) enabled us to estimate the lateral width in the 3D confined structures. We also investigated the optical properties of the grown structures by low-temperature micro-photoluminescence (μ -PL) measurements.

II. FABRICATION OF CONFINED STRUCTURES

GaAs (001) substrates were patterned using electron-beam lithography followed by wet chemical etching. The negative-type resist was SAL601 (Shipley) and etchant was $\text{H}_3\text{PO}_4:\text{H}_2\text{O}_2:\text{H}_2\text{O}$ (1:2:40) at 24 °C. The pattern comprised square mesa structures with sides oriented along $\langle 100 \rangle$ directions. The square mesas had lateral widths of 2.8–4.2 μm with an interval of 0.2 μm after etching. The size variation allowed us to fabricate single 3D structures of various lateral widths on isolated pyramidal structures in the same growth run.

Epitaxial growth was performed in a conventional MBE system. The growth run was done twice with different thicknesses of the GaAs buffer layer (sample A: 500 nm and sample B: 700 nm). The growth structures consisted of the

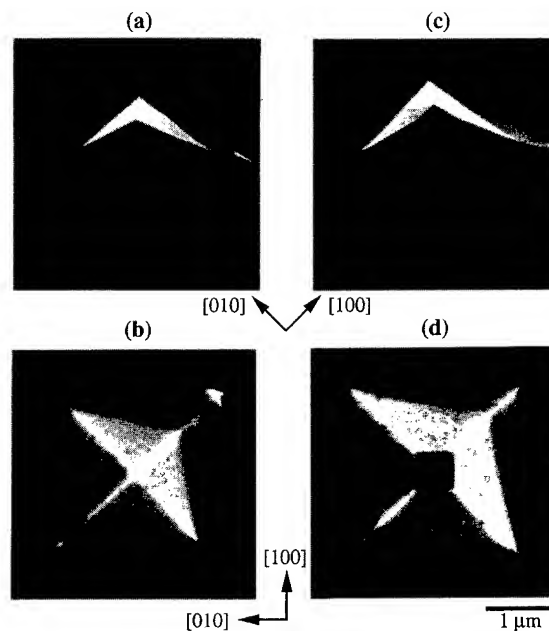


FIG. 1. SEM images of pyramidal structures grown on square mesas in sample B. (a) and (b) Images of a pinched off pyramidal structure after growth on the 2.8 μm mesa. The pyramid is limited by sidewall $\{011\}$ facets. (c) and (d) Images of a truncated pyramidal structure after growth on the 3.2 μm mesa. The pyramid is limited by sidewall $\{011\}$ facets and a mesa-top facet. The lower images, (b) and (d), are top views of the same pyramidal structures as in the upper images.

GaAs buffer layer, AlAs/GaAs/AlAs QW layers (10 nm/5 nm/10 nm), and a GaAs capping layer (10 nm). The layer thickness was estimated from the growth time of each layer and from the growth rates of GaAs and AlAs on a flat GaAs (001) substrate, which were both 0.5 $\mu\text{m}/\text{h}$. The growth temperature was 660 °C, which was monitored by a thermocouple placed on the rear of the substrates. Arsenic beam equivalent pressure was 7×10^{-6} Torr, and the substrate rotation during growth was 20 rpm. For structural and optical characterizations sample B was used in the following sections except in Sec. IV A, where sample A was used.

III. STRUCTURAL CHARACTERIZATION

A. Scanning electron microscopy

Figure 1 shows SEM images of typical structures grown on 2.8 μm side (left) and 3.2 μm side (right) square mesas in sample B. The images in the lower half of Fig. 1 are top views of the same pyramidal structure as shown in the upper half. The grown structures were pinched-off and truncated pyramids limited by sidewall $\{011\}$ facets and the mesa-top facet.

As shown in Fig. 1, the sidewalls had relatively rough surface morphology. The roughness increased when the same pyramidal structures were observed again a few months later. This suggests that the roughness is due to oxidation of the AlAs barrier layers resulting from the thinness of the GaAs capping layer on the sidewalls. In contrast to the sidewalls, the mesa top of the truncated pyramidal structure had smooth surface morphology, as shown in Fig. 1(d). The

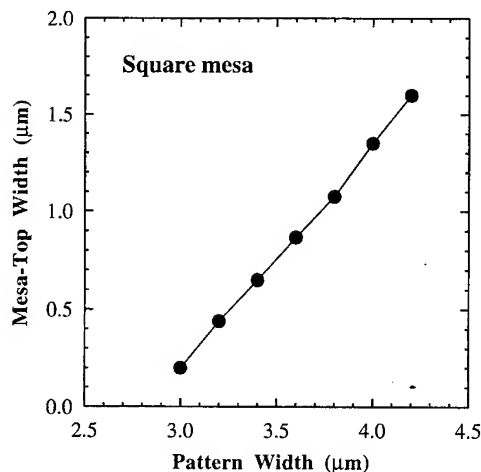


FIG. 2. Mesa-top width of pyramidal structures estimated by SEM observation as a function of pattern width. The mesa-top width reflects the lateral width of the 3D confined structure in the pyramid.

GaAs capping layer on the mesa top was expected to be thicker than that on the sidewalls because of the faster growth rate of GaAs on the mesa top. This greater thickness may prevent the oxidation of the AlAs layers on the mesa top.

The sidewalls connected with the bottom regions of the patterned substrates during growth of the GaAs buffer layer in the present study. As shown in Fig. 1, accumulation of GaAs at the base of the pyramidal structures differed along the $[110]$ and $[1\bar{1}0]$ directions. The pyramidal structures were more buried along the $[1\bar{1}0]$ direction. Owing to this anisotropy, the shape of the pyramidal structures is a rhombus in the top-view images of Figs. 1(b) and 1(d). López *et al.* suggested¹⁰ that the anisotropy could be due to the different lateral growth rates¹⁴ along the $[110]$ and $[1\bar{1}0]$ directions.

The mesa-top shape of the truncated pyramidal structure in Fig. 1(d) was square and equivalent to that of the as-patterned mesas. This mesa-top shape reflects the lateral feature of the 3D confined structure embedded in the pyramidal structure. As mentioned in Sec. I, the confining potential wells in the lateral directions were due to the lateral variations in the QW layer thickness of the mesa top and the sidewalls. The shapes of the lateral potential wells are thus determined by the mesa-top shapes just before growth of QW layers. Moreover, the mesa-top shapes do not change rapidly during growth of the QW heterostructure and the capping layer if their layer thickness is much less than the mesa-top width. The lateral width of the 3D confined structures can therefore be estimated from the mesa-top width after growth.

Figure 2 shows the dependence of the mesa-top width on the pattern width, except for the smallest one whose pyramidal structure was pinched off according to the SEM observation. As shown in Fig. 2, the mesa-top width, that is the lateral width of the 3D confined structure, decreased from 1.6 to 0.2 μm in proportion to the decrease in pattern width.

This conflicted with the expectation that a strong increase in growth rate on the mesa top would promote a rapid reduction in the lateral width.^{8,15} However, the suppression of the rapid reduction in the lateral size can be explained by adatom migration on the sidewalls towards the bottom region in the present case. If an almost equal number of adatoms migrate from the sidewalls to the mesa top, the shrinkage of the mesa-top area makes the growth rate increase. However, if the sidewalls connect with the bottom, some of the adatoms on the sidewalls migrate to the bottom, and consequently fewer adatoms migrate to the mesa top. This decrease may suppress the increase in the growth rate and may keep the proportional relationship between the lateral width and the pattern width.

Konkar *et al.* have reported that the mesa top of pyramidal structures experienced two stages of pinch off related to the initial formation of sidewall $\{011\}$ facets followed by the formation of $\{013\}$ facets.⁹ From TEM observation, they found that the $\{013\}$ facets caused an early pinch off and $\{011\}$ facets subsequently took over and caused a second pinch off. In the present study, however, the pinch off occurred in only one stage related to the formation of sidewall $\{011\}$ facets. This discrepancy may arise from a difference in the growth conditions, such as the growth temperature, which changes the migration rates of adatoms on neighboring facets.

B. Atomic force microscopy

Although the SEM observation indicated that the pyramidal structure on the smallest mesa pattern was pinched off and had no flat top, it could have missed a flat region on the top because of inadequate spatial resolution. We therefore observed the top shape of the pyramidal structure on the smallest mesa pattern with an AFM, which has extremely high spatial resolution. The cross-sectional surface profiles of the pyramidal structure in the $[100]$ and $[010]$ directions [Fig. 3(a)] do not show a flat region on the top. However, they showed the top shape to be round rather than pointed.

A round top was also previously reported on pinched off pyramidal structures grown by MOCVD on masked GaAs (001) substrates with square openings.¹⁶ That report suggested that the top shape of the pinched off pyramidal structures was determined by self-limiting growth depending on growth conditions. In our case, the extent of the round region was about 30 nm, which was close to the diameter of the probe tip in the AFM. Thus we believe that the round top of the pyramidal structure in the present case is affected by the shape of the probe tip.

The AFM was also applied to the truncated pyramidal structures in order to observe their flatness. Figure 3(b) shows cross-sectional surface profiles of the truncated pyramidal structures in the $[100]$ and $[010]$ directions. These profiles revealed that the mesa-top plane tilted to the $[100]$ direction by 10° and that the edges of the top plane rose higher than the center. The rise can be explained by a good supply of adatoms due to efficient migration from the sidewalls to the mesa top. The tilting of the mesa-top plane may suggest

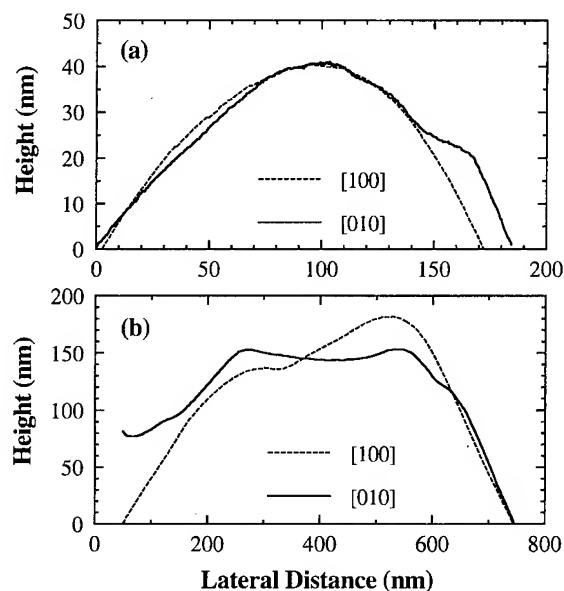


FIG. 3. Cross-sectional surface profiles of (a) a pinched off pyramid and (b) a truncated pyramid obtained by AFM observation. The profiles in the [100] and [010] directions are shown by broken and solid lines, respectively. Note that the scales of the axes are different in the figures.

that under the present growth conditions the migration rate of the adatoms differs among the sidewall {011} facets surrounding the top plane.

IV. OPTICAL CHARACTERIZATION

We performed spectroscopy on single 3D confined structures by positioning and holding the mesa-top region of one isolated pyramidal structure at a time under laser illumination. The photoluminescence (PL) and photoluminescence excitation (PLE) spectra were obtained by using a low-temperature μ -PL method that uses confocal microscopy,¹⁷ a form of scanning optical microscopy with far-field optics. High spatial resolution is attained by scanning a focused spot of excitation light on an object in the confocal arrangement, which has an aperture on the image focal plane in conjunction with the excitation spot.

The optical setup for the μ -PL measurements was the same as in our previous reports.¹⁸ A laser beam was focused and luminescent light was collected by a microscope objective through the optical window of a He-flow cryostat. The objective was specially designed to eliminate any spherical aberration caused by window plates in the He cryostat and has a magnification of 45 and a numerical aperture of 0.5. This objective focused a laser beam of 514.5 nm excitation wavelength into a 1.5 μ m-diam spot on the sample surface in the He cryostat. Details of the optical setup are given elsewhere.¹⁸

In the present study, the aperture in the confocal arrangement was a 50 μ m pinhole, which gave a spatial resolution of less than 0.7 μ m. The excitation sources were an Ar-ion laser and a tunable Ti:Sapphire laser. All measurements were done at a sample temperature of 8 K.

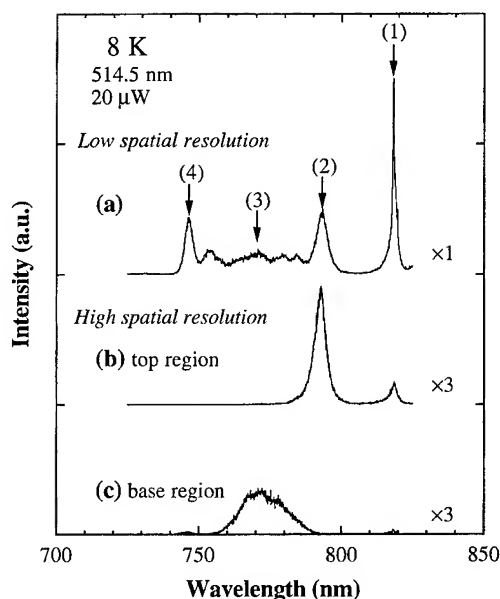


FIG. 4. PL spectra of a truncated pyramidal structure. Excitation wavelength was 514.5 nm and excitation power was 20 μ W. (a) PL spectra of the pyramidal structure obtained by measurement with low spatial resolution. This PL spectrum exhibits some emission peaks due to various QW layers included in the pyramidal structure. Arrows point to detection wavelengths in the measurements of the PL images shown in Fig. 5. (b) and (c) PL spectra of the pyramidal structure obtained by measurements illuminating only (b) its mesa-top region or (c) the base region.

A. Photoluminescence images of a pyramidal structure

We first measured PL spectra of a truncated pyramidal structure in sample A with low spatial resolution, which was attained by defocusing the laser beam to a 6- μ m-diam spot on the sample surface and removing the pinhole aperture. This measurement gave the PL spectra of various QW layers included in the pyramidal structure. The resultant spectrum, shown in Fig. 4(a), exhibits some distinct peaks. To attribute the spatial origins of these PL peaks, we next measured the PL images with high spatial resolution. The PL images were obtained by detecting luminescent light while scanning the focused spot of excitation light across the pyramidal structure.

Figure 5(a) shows spectrally and spatially resolved PL images of the pyramidal structure with detection wavelength indicated by arrows in Fig. 4(a), and Fig. 5(b) shows a SEM image of its top view. The PL images reveal that the peak at 792.8 nm (2) is due to luminescence from the single 3D confined structure on the mesa top of the pyramidal structure, whereas the peak at 746.5 nm (4) corresponds to emissions from the QW heterostructure at the bottom region between mesas patterned on the substrate. The attribution of the peak (4) to the mesa-top region was verified by a PL measurement illuminating only the mesa-top region, as shown in Fig. 4(b). The PL images also reveal that the longest wavelength peak of the spectrum in Fig. 4(a) at 818.1 nm (1) originates from the GaAs buffer layer, and that the broad peak at 770.0 nm (3) originates from the QW heterostructure

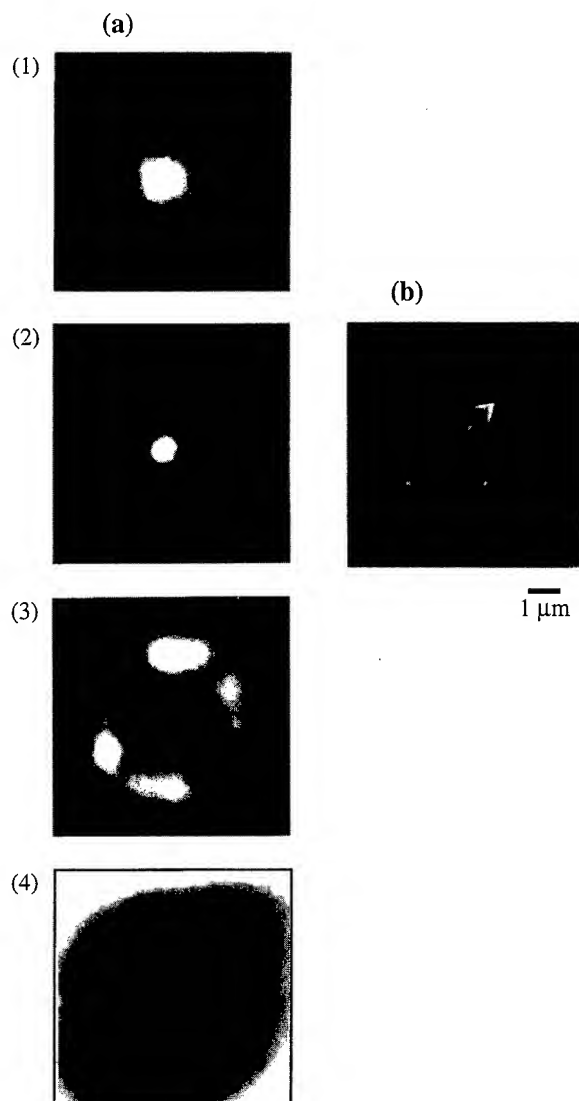


FIG. 5. (a) PL images of the truncated pyramidal structure used in the measurements of the PL spectra shown in Fig. 4. The images were obtained by scanning a focused spot of excitation light across the pyramidal structure. The scanning area of each PL image was $8\ \mu\text{m} \times 8\ \mu\text{m}$ with a scanning interval of $0.2\ \mu\text{m}$. The number of pixels in the PL image was increased by interpolation. Excitation power was $20\ \mu\text{W}$ and the detection wavelengths in the PL image were (1) 818.1 nm, (2) 792.8 nm, (3) 770.7 nm, or (4) 746.5 nm, which correspond to the arrows in Fig. 4(a). (b) SEM image of the truncated pyramidal structure.

around the base region of the pyramidal structure.

Since the thickness of a QW layer is estimated from its emission wavelength, we can determine the QW layer thickness at each region mentioned above. At the bottom region its thickness was estimated to be about 5 nm, which equals the design value. This indicates that the mesa patterns had little effect on the growth rates at the well-lighted bottom region in Fig. 5(a) (4).

The mesa-top width of the measured pyramidal structure was estimated to be about $0.9\ \mu\text{m}$ from the SEM image in Fig. 5(b). The 3D confined structure in the pyramidal structure was thus regarded as a conventional QW heterostructure. Accordingly, the GaAs thickness of the 3D confined

structure can also be estimated from its emission wavelength. The estimated value was about 9 nm, which is much larger than that at the bottom region. This large value indicates efficient adatom migration from the sidewalls to the mesa top.

Efficient migration was also expected to make the QW layer thinner at the sidewall regions than at the bottom regions. Since we could not attribute any peak to the QW heterostructure at the sidewall regions in the PL spectrum of Fig. 4(a), the QW layer must be much thinner at the sidewall regions than at the bottom and mesa-top regions.

The PL image revealed that luminescence from the QW heterostructure around the base region corresponds to the peak at 770.0 nm. However, the QW layer thickness is not estimated only by this wavelength because of the broad spectral width. When we illuminated only the base region of the pyramidal structure, the PL spectrum exhibited a cluster of broad peaks located between two peaks due to the emissions from the mesa-top and bottom regions, as shown in Fig. 4(c). Thus the QW layer thickness at the base region varied widely and lay between those at the mesa-top and bottom regions.

B. Photoluminescence spectra of single confined structures

Figure 6 shows the typical PL spectra of single 3D confined structure of various lateral widths obtained by measuring the isolated pyramidal structures in sample B. All the spectra except for the pinched off pyramidal structure exhibit distinct peaks due to exciton luminescence from the 3D confined structures. Although the spectrum of the pinched off pyramidal structure exhibits some peaks, the measurement of PL images revealed that there was no spectral peak corresponding to the emissions from the mesa top.

The SEM and AFM observations mentioned in Secs. III A and III B found that the pyramidal structure on the smallest mesa pattern was pinched off and did not have a flat top. If a 3D confined structure were formed in this pinched off pyramidal structure, its size would be expected to be so small that three-dimensional quantum effects would be evident. The three-dimensional quantum effects shift the PL peak to the shorter wavelength side compared to a conventional QW having the same layer thickness. In the present case, the shifted peak would be located within a wavelength range where the luminescence from the base region is strong, and consequently the PL peak of the 3D confined structure would be obscured by the peaks due to luminescence from the base region unless luminescence from 3D confined structure were very strong. Thus we cannot determine from the PL measurements whether or not the 3D confined structure was formed in the pinched off pyramidal structure.

Since the minimum lateral width of the 3D confined structures in the present study was about $0.2\ \mu\text{m}$, the emission light from them was regarded as that from conventional QW heterostructures. We can thus investigate the dependence of the vertical layer thickness in the 3D confined structures on the lateral width by means of their emission wavelength. In

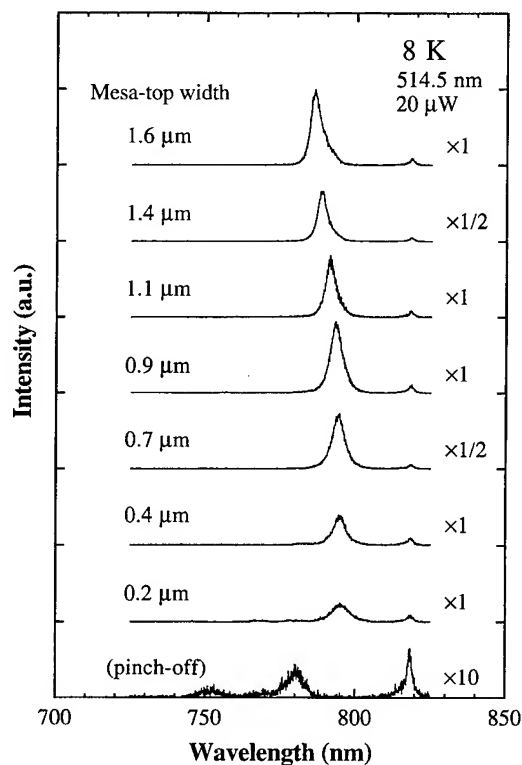


FIG. 6. PL spectra of isolated pyramidal structures of various sizes illuminating only their mesa-top region. Excitation wavelength was 514.5 nm and excitation power was 20 μ W. All the spectra except the pinched off pyramidal structure exhibit distinct peaks due to the exciton luminescence from single 3D confined structures. Although the spectrum of the pinched off pyramidal structure exhibits some peaks, there is no peak corresponding to the emissions from its mesa top.

Fig. 7, the peak wavelength is plotted (filled symbols) as a function of the lateral width of the 3D confined structures. The emission peak wavelength shifted to the longer side as the lateral width decreased. This shift implies that the QW layer thickness, that is the GaAs growth rate, increased as the

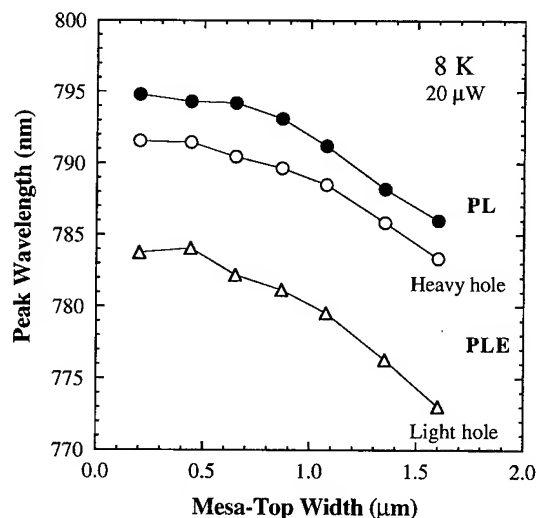


FIG. 7. Peak wavelengths in PL (filled symbols) and PLE (open symbols) spectra as a function of lateral width in 3D confined structures.

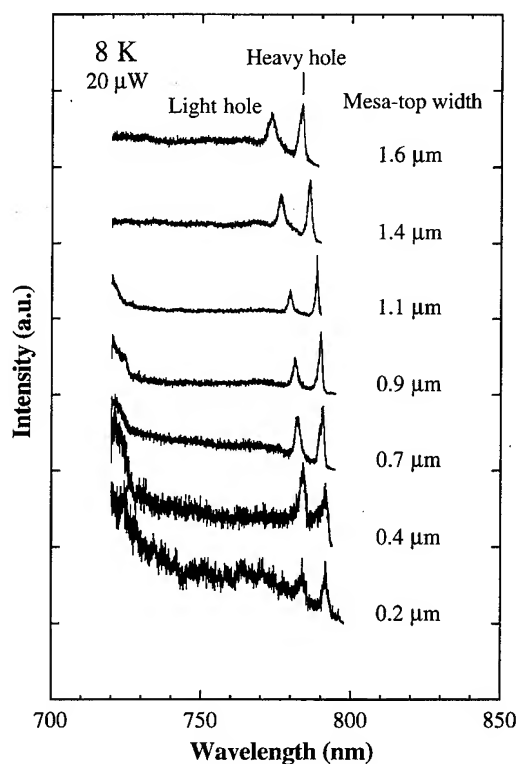


FIG. 8. PLE spectra of single 3D confined structures of various lateral widths. The confined structures were the same as the ones used in the PL measurements. Detection wavelength was located on the peak position of each PL spectrum in Fig. 6 and excitation power was 20 μ W. Each PLE spectrum exhibits two distinct peaks, which arise from the recombination of heavy-hole and light-hole excitons.

lateral width decreased. We estimate the layer thickness increased from 8 to 10 nm. However, this increase saturated at smaller lateral widths. As mentioned in Sec. III A, some reports indicated that the growth rate on the mesa top strongly increased as the lateral size decreased.^{8,15} The saturation in the present study may also be explained by adatom migration on the sidewalls towards the bottom region.

C. Photoluminescence excitation spectra of single confined structures

Typical PLE spectra of the single 3D confined structures of various lateral widths are shown in Fig. 8. The confined structures were the same as the ones used in the PL measurements. Each PLE spectrum exhibits distinct two peaks, which arise from the recombination of heavy-hole and light-hole excitons. Such distinct exciton peaks have been never shown in PLE spectra of 2D or 3D confined structures grown on patterned substrates.⁴ The reason that distinct peaks were obtained in the present study may be related to the buffer layer being thick, more than 1 μ m.

In Fig. 7, the peak wavelength of the PLE spectra is plotted (open symbols) as a function of lateral width in the 3D confined structures. Like the PL spectra, the exciton peaks in the PLE spectra shifted to longer than a wavelength with

decreasing lateral width. The shift also saturated at smaller lateral widths. In addition, the peaks in the PL spectra were always at longer wavelengths than those in the PLE spectra.

The PL peaks are due to the emissions from the recombination of heavy-hole excitons. Figure 7 thus indicates that the heavy-hole-exciton peak in the PL spectra shifted to a longer wavelength in the PLE spectra. This red shift of the exciton peak is often seen in conventional QW heterostructures, where it is explained by a localization effect of the excitons. The peaks in the PLE spectra are due to absorption creating free excitons, while those in the PL spectra are due to the recombination of the excitons trapped by the QW layer thickness fluctuation. However, as will be discussed in the next section, the red shift in the present case is not entirely due to the localization effect, but also depends on the broad distribution in QW layer thickness.

D. Excitation-power dependence of spectra

The PL spectra of the single 3D confined structures shown in Fig. 6 exhibit a broad peak about 10 meV wide. Since the structures were single, the broad width cannot be explained by size fluctuation. Another possible origin of the peak broadening may be a band filling effect. If bands are filled with carriers to higher energy sites, the carriers will also recombine with higher luminescent energy, and then the luminescent peak will broaden. In the case of 3D confined structures, some carriers generated by laser illumination are trapped in a small volume, creating an extremely high carrier density in these structures, even if low excitation power is employed. In addition, the excitation power of 20 μW in Fig. 6 is already very large since the laser spot diameter is on the micrometer order. If the carriers generated by this excitation power are effectively trapped in a small volume such as 3D confined structures with μm -order lateral size, the carrier density becomes extremely high, and the band filling effect will occur. To clarify the origin of spectral broadening, we thus measured the dependence of the PL and PLE spectra on excitation power.

Figure 9 shows the PL spectra of the 3D confined structure of the 1.4 μm lateral width with various excitation powers. The spectrum with 200 μW excitation power (top) exhibits a broad peak about 20 meV wide. As the excitation power is reduced, however, the spectra became a cluster of sharp peaks. Each peak in the cluster may be due to the emissions from the recombination of the excitons trapped in lower-energy sites of energy fluctuation in the QW heterostructure,¹⁹⁻²¹ and may be sharpened by reduction of the band filling effect. However, broad distribution of the sharp peaks remained in the PL spectra even with the lowest excitation power. This suggests that the broad spectral distribution is caused mainly by wide variation in QW layer thickness.

Figure 10 shows the PLE spectra of the 3D confined structure of the 1.4 μm lateral width for various excitation powers, including its PL spectrum with a 720 nm excitation wavelength and a 2 μW excitation power. As the excitation power decreased, the light-hole-exciton peaks varied little,

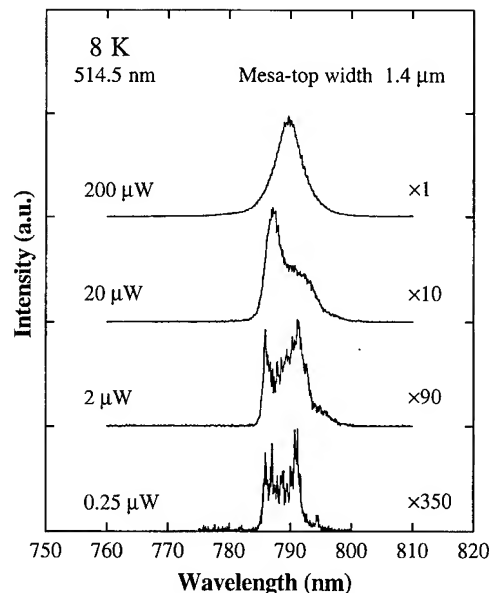


FIG. 9. PL spectra of single 3D confined structure of 1.4 μm lateral width with various excitation powers. Excitation wavelength was 514.5 nm.

but the heavy-hole-exciton peaks did vary slightly. This variation in the heavy-hole-exciton peaks may be related to the variation in the PL peaks depending on excitation powers. However, we do not know why only the peaks of heavy-hole excitons are affected by excitation power. Further study is required to clarify the excitation-power dependence of the PLE spectra.

As mentioned above, the heavy-hole-exciton peaks in the PL spectra shifted to a longer wavelength than those in the

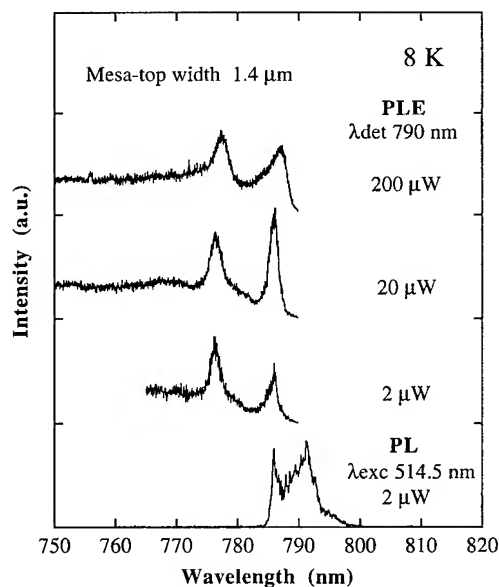


FIG. 10. PLE spectra of single 3D confined structure of 1.4 μm lateral width with various excitation powers and the PL spectrum with 514.5 nm excitation wavelength and 2 μW excitation power. Detection wavelength in PLE spectra was 790 nm.

PLE spectra. In the case of low excitation power shown in Fig. 10, however, a peak in the PL cluster located at the peak position in the PLE spectra. The red shift in the exciton peaks shown in Fig. 7 is thus dependent on excitation power. The broad distribution of QW layer thickness may be the main cause of the red shift in the spectra with high excitation power.

V. CONCLUSION

We have reported structural and optical characteristics of single 3D confined structures grown on square mesas patterned onto GaAs (001) substrates. By fabricating single structures and using AlAs barrier layers, we were able to clearly attribute luminescent spectral peaks to spatial origins in the grown structures.

We first observed surfaces of the grown structures with a SEM and an AFM. The observation indicated that the structures on the mesas were pinched off and truncated pyramids limited by sidewall {011} facets and a mesa-top facet. The observation led to estimation of lateral width in the 3D confined structures, which decreased from 1.6 to 0.2 μm in proportion to a decrease in pattern width. This proportionality conflicts with the expectation that a strong increase in growth rate on the mesa top promotes a rapid reduction in their lateral widths. However, the suppression of the rapid reduction in the present study can be explained by adatom migration on the sidewalls towards the bottom region of the patterned substrates.

We next investigated the optical properties of the grown structures by $\mu\text{-PL}$ measurements at 8 K. The PL images were measured in order to attribute luminescent spectral peaks to the QW heterostructures spatially originating in the pyramidal structures. When estimated by using the luminescent wavelength, the GaAs layer thickness in the 3D confined structure was rather larger than that at the bottom. In addition, the PL images showed no spectral peak identified with the QW heterostructure at the sidewalls. This means that the QW layer at the sidewalls is rather thinner than that at the bottom and the mesa top.

We performed spectroscopy on single 3D confined structures of various widths by positioning and holding the mesa top of one isolated pyramidal structure at a time under laser illumination. The PL spectra exhibited distinct exciton peaks about 10 meV wide. As the lateral width decreased, the vertical layer thickness estimated by the emission wavelength increased from 8 to 10 nm, but saturated at smaller lateral widths. The saturation can also be explained by adatom migration on the sidewalls towards the bottom region. The PLE spectra also showed distinct exciton peaks. The reason that distinct peaks were obtained in the present study may be related to the buffer layer being thick, more than 1 μm .

The broad peak width in the PL spectra cannot be due to size fluctuation because we measured single structures. When the excitation power was reduced, PL spectra became a cluster of sharp peaks. Each peak in the cluster may be due

to the recombination of the excitons trapped in QW layer thickness fluctuation and may be sharpened by reduction of the band filling effect. The cluster thus suggests that the broad spectral width is caused mainly by wide variation in QW layer thickness in the 3D confined structures.

By performing growth on patterned substrates, we fabricated 3D confined structures on a sub- μm scale that had high optical quality. However, we had some problems controlling the formation of nanostructures such as quantum dots using this growth technique. First, it was difficult to control lateral size of 3D confined structures on a nanometer scale, since fluctuation in pattern size caused a fluctuation in the lateral size owing to the relationship between them. There was also a wide variation in QW layer thickness of 3D confined structures. This may be caused by the difference in growth conditions preferred for growth of QW heterostructures and growth of the pyramidal structure itself. These problems suggest that another growth mechanism is required for fabricating nanostructures on top of pyramids.

ACKNOWLEDGMENT

The authors would like to thank Dr. M. López for his valuable discussion of growth on patterned substrates.

¹*Proceedings of 22nd International Conference on the Physics of Semiconductors*, Vancouver, Canada, 1994, edited by D. J. Lockwood (World Scientific, Singapore, 1995).

²*Proceedings of the Seventh International Conference on Modulated Semiconductor Structures*, Madrid, Spain, 1995, edited by C. Tejedor and L. Viña [Solid-State Electron. **40**, 1 (1996)].

³*Nanostructures and Quantum Effects*, edited by H. Sakaki and H. Noge (Springer, Berlin, 1994).

⁴*Semiconductors and Semimetals*, edited by A. C. Gossard (Academic, New York, 1994), Vol. 40.

⁵A. Madhukar, K. C. Rajkumar, and P. Chen, Appl. Phys. Lett. **62**, 1547 (1993).

⁶K. C. Rajkumar, A. Madhukar, K. Rammohan, D. H. Rich, P. Chen, and L. Chen, Appl. Phys. Lett. **63**, 2905 (1993).

⁷M. López, T. Ishikawa, and Y. Nomura, Electron. Lett. **29**, 2225 (1993).

⁸K. C. Rajkumar, A. Madhukar, P. Chen, A. Konkar, L. Chen, K. Rammohan, and D. H. Rich, J. Vac. Sci. Technol. B **12**, 1071 (1994).

⁹A. Konkar, K. C. Rajkumar, Q. Xie, P. Chen, A. Madhukar, H. T. Lin, and D. H. Rich, J. Cryst. Growth **150**, 311 (1995).

¹⁰M. López, N. Tanaka, I. Matsuyama, and T. Ishikawa, Appl. Phys. Lett. **68**, 658 (1996).

¹¹D. H. Rich, H. T. Lin, A. Konkar, P. Chen, and A. Madhukar, Appl. Phys. Lett. **69**, 665 (1996).

¹²H. P. Meier, E. Van Gieson, W. Walter, C. Harder, M. Krahle, and D. Bimberg, Appl. Phys. Lett. **54**, 433 (1989).

¹³I. Matsuyama, M. López, N. Tanaka, and T. Ishikawa, Jpn. J. Appl. Phys. **33**, L627 (1994).

¹⁴M. Kawabe and T. Sugaya, Jpn. J. Appl. Phys. **28**, L1077 (1989).

¹⁵Y. D. Galeuchet, P. Roentgen, and V. Graf, J. Appl. Phys. **68**, 560 (1990).

¹⁶K. Kumakura, K. Nakakoshi, J. Motohisa, T. Fukui, and H. Hasegawa, Jpn. J. Appl. Phys. **34**, 4387 (1995).

¹⁷*Confocal Microscopy*, edited by T. Wilson (Academic, London, 1990).

¹⁸J. Kasai and Y. Katayama, Rev. Sci. Instrum. **66**, 3738 (1995); **67**, 4397 (1996).

¹⁹H. F. Hess, E. Betzig, T. D. Harris, L. N. Pfeiffer, and K. W. West, Science **264**, 1740 (1994).

²⁰K. Brunner, G. Abstreiter, G. Böhm, G. Tränkle, and G. Weimann, Appl. Phys. Lett. **64**, 3320 (1994).

²¹D. Gammon, E. S. Snow, B. V. Shanabrook, D. S. Katzer, and D. Park, Phys. Rev. Lett. **76**, 3005 (1996).

Electrical transport properties of silicon delta-doped $\text{Al}_{0.30}\text{Ga}_{0.70}\text{As}$ samples showing suppression of the DX center features

J. A. Corrêa F.,^{a)} A. G. de Oliveira,^{b)} M. I. N. da Silva, G. M. Ribeiro, and J. F. Sampaio
*Departamento de Física, Instituto de Ciências Exatas, Universidade Federal de Minas Gerais,
31161-970 Belo Horizonte, Minas Gerais, Brazil*

(Received 24 January 1997; accepted 30 May 1997)

Photoconductivity and photo-Hall density measurements using an infrared light emitting diode as the light source were carried out on single silicon delta-doped $\text{Al}_{0.30}\text{Ga}_{0.70}\text{As}$ samples as a function of temperature. The samples were grown by molecular beam epitaxy at 530 °C and 600 °C. We have studied the effect of etching the cap layer on the electrical transport properties. An observed persistent photoconductivity effect is explained using a model of parallel conduction in two nearby spatially separated channels. We will present evidence that the DX center is not active for nearly ideal delta-doped samples. We have proposed that the DX-center level related to the conduction-band minimum is strongly dependent on the silicon delta-doping density and on the growth conditions. © 1997 American Vacuum Society. [S0734-211X(97)15304-5]

I. INTRODUCTION

With the development of epitaxial growth techniques high frequency electronic devices based on AlGaAs/GaAs heterostructures have been obtained. The essence of most devices is the presence of the AlGaAs/GaAs interface near a doped AlGaAs layer. This structure, usually called the modulation doping structure (MDS), produces a quasi-two-dimensional electron gas (2DEG) of high mobility formed at the interface. However, the presence of DX centers in the silicon uniformly doped AlGaAs layer causes instabilities in device operation due to a persistent photoconductivity (PPC) effect. In order to reduce, at least partially, the density of silicon atoms acting as DX centers replacement of the uniformly doped AlGaAs layer by a silicon delta-doped layer was proposed.¹ Most studies on silicon delta-doped AlGaAs samples have been carried out in multi-delta-doped layers and their main aim was to study the interaction between the various delta-doped wells. In order to assess the consequence of the replacement, it is advisable to distinguish those effects arising (a) from the delta-doped layer itself, (b) from the heterojunction, and (c) from the coupling between them. Very little effort has been made in studying single delta-doped AlGaAs samples, and that done has mainly used the Hall effect technique; the reason, as suggested by some authors,^{2,3} involves difficulties concerning DX centers. The few measurements on single delta-doped samples include capacitance-voltage techniques (*C-V*) and secondary ion mass spectroscopy (SIMS), but their aim was primarily the study of the doping profile on samples grown by different techniques.⁴⁻⁶ The two-dimensional (2D) features of DX centers have been studied by photoluminescence⁷ and Hall effect techniques^{1,8} in multi-delta-doped samples, and in a less systematic way on single delta-doped samples by *C-V*⁹

and magnetoresistance.¹⁰ The features of the DX center are not yet firmly established for delta-doped structures. Solomon *et al.*,⁸ using deep level transient spectroscopy (DLTS), observed an enhancement of about six orders of magnitude in the capture cross section of DX centers in two dimensions, and a reduction of about 40% in the silicon atoms acting as DX centers by Hall effect measurements. The DX-center level was found to be 60 meV below the conduction-band edge, i.e., for the used aluminum content it was 40 meV shallower than in a uniformly doped sample with an equivalent 2D silicon concentration. The authors have suggested that the 2D features of DX center are affected by a biaxial stress state, which is induced by the large silicon concentration rather than by the 2D confinement. Etienne and Thierry-Mieg¹ suggested something similar when they invoked a nonhomogeneous stress state due to a high silicon concentration confined in a narrow region of the sample. However, Arscott *et al.*,¹¹ using *C-V* and DLTS techniques, reported the observation of a PPC effect in a delta-doped GaAs sample with a very high silicon concentration ($\approx 10^{13} \text{ cm}^{-2}$). This is somehow unexpected since for GaAs the DX-center level is resonant with the conduction band. Mejri *et al.*⁷ reported that DX-center features were suppressed in heavily (or lightly) silicon delta-doped $\text{Al}_{0.32}\text{Ga}_{0.68}\text{As}$ samples when the large (or short) periodical delta-doped structure exhibited superlattice properties. According to Bourgoin *et al.*,¹² the suppression of DX-center features occurs when the original structure of the *L* band is not preserved and not due to stress. Nevertheless, Chadi and Chang¹³ showed that the binding energy of the DX center corresponds to an average over the conduction bands (Γ , *L* and *X*) and not only to the *L* band. The structural model assumes that the potential of the impurity in a three-dimensional (3D) uniform distribution is not strong enough to cause a mixing of the conduction bands. This should, however, be considered in delta-doped structures due to the

^{a)}On leave from the Departamento de Ciências Naturais, Fundação de Ensino Superior de São João Del Rei, Caixa Postal 110, 36300-000 São João del Rei, Minas Gerais, Brazil.

^{b)}Electronic mail: alfredo@oraculo.lcc.ufmg.br

high electron field formed in the well region. Summarizing, 2D features of the DX center are far from being firmly established.

The study of 2D transport features of silicon delta-doped AlGaAs samples demands both low silicon concentrations and low growth temperatures in order to avoid self-compensation, segregation and diffusion effects.^{3,14} The two last effects broaden the silicon profile, thus destroying the features that might be studied. On the other hand, samples with low silicon concentration are highly resistive. We believe that this is the reason why multi-delta-doped rather than single delta-doped samples are usually being studied in order to assess the 2D features. For samples with multi-delta-doped layers the electrical signal is reinforced, and frequently the AlGaAs layer is uniformly and lightly doped.⁸ Finally, it should be noted that most of the results reported in the literature refer to samples grown at temperatures considered optimal for electrical and optical measurements, namely 600–700 °C. At these temperatures the 3D features of DX centers are stressed.

In the present work we report on the electrical properties of single silicon delta-doped Al_{0.30}Ga_{0.70}As samples with emphasis on the 2D features of DX centers studied by means of the Hall effect. In an attempt to reduce the spread through diffusion we have grown the samples at 530 °C and to enhance the concentration of the DX centers the silicon concentration used was higher than $8 \times 10^{12} \text{ cm}^{-2}$. We have observed that, under conditions of darkness and below 150 K, the Hall density increases for decreasing temperature. This behavior is exactly the opposite of what occurs for uniformly doped samples (see, for example, Ref. 1). We have also observed that by cooling down some of the samples under illumination their electrical characteristics present a transition from *n* to *p* type. Finally, the DX center presents a very low efficiency, seeming to contribute little to defining the electrical properties of the samples. We have proposed that the DX-center level must be resonant with the conduction band for nearly ideal delta-doped Al_{0.30}Ga_{0.70}As samples.

II. EXPERIMENT

The single silicon delta-doped Al_{0.30}Ga_{0.70}As samples were grown by molecular beam epitaxy (MBE) on a semi-insulating GaAs(100) substrate. A 0.5 μm thick unintentionally doped GaAs buffer layer was grown at 600 °C. Then, without interrupting the growth process, the temperature was reduced to 530 °C, and was kept constant for the remaining growth steps. After temperature adjustment, a 1 μm thick unintentionally doped Al_{0.30}Ga_{0.70}As layer was grown (to prevent a remote 2D gas at the GaAs/Al_{0.30}Ga_{0.70}As interface), and growth was interrupted by closing the gallium and aluminum shutters. Next, after a pause of 1 min in order to smooth the surface, the silicon shutter was opened for the time necessary to achieve the desired silicon concentrations of 8.4, 12.6 and $20 \times 10^{12} \text{ cm}^{-2}$ for sample Nos. 236, 233 and 239, respectively. One minute after the silicon deposit, the gallium and aluminum shutters were opened in order to grow

a 0.5 μm thick unintentionally doped Al_{0.30}Ga_{0.70}As layer. Finally, all samples were covered with a 170 Å *n*⁺-GaAs ($4 \times 10^{18} \text{ cm}^{-3}$) cap layer. A fourth sample, No. 231, was grown under the same conditions as the sample No. 236, except that the growth temperature was 600 °C. The aluminum content of $x=0.30$ was chosen in order to enhance the DX center features.

The devices were photolithographically patterned and etched into Hall bars. The ohmic contacts were made by annealing In-Sn (In-Zn for sample No. 231) at 400 °C in a N₂-H₂ (85:15) atmosphere for 10 min. The Hall effect measurements were performed carefully in darkness and under illumination using an infrared light emitting diode (LED) of GaAs (1.32 eV at room temperature) in a closed cycle helium gas cryostat in the temperature range from 4 K to room temperature under magnetic field intensity of 0.7 T. All measurements were carried out using a cooling or heating rate of 2 K/min. The Hall density and mobility were obtained after an interpolation of the data for accuracy of the results as suggested by Brunthaler *et al.*¹⁵ The current-voltage (*I*-*V*) characteristic has shown a linear behavior for *I_S* below 5 mA. The measurements were carried out using *I_S* = 20 μA .

III. RESULTS AND DISCUSSION

We have obtained considerable experimental evidence suggesting that the DX center is not present at all or, at least, is not active in single silicon delta-doped AlGaAs alloys. The experimental results as well as our discussion about them as follows.

A. Absence of an Arrhenius adjustment

Figure 1 shows the Hall density n_{Hall} as a function of temperature reversal in darkness to cool down the samples. At room temperature, the Hall density value of $8.3 \times 10^{12} \text{ cm}^{-2}$ for sample No. 236 (downward triangle) and $11.7 \times 10^{12} \text{ cm}^{-2}$ for No. 233 (diamond) compares well with the silicon concentration, suggesting a low self-compensation effect. On the other hand, the Hall density for sample No. 239 (upward triangle) was found to be $3.2 \times 10^{12} \text{ cm}^{-2}$, which suggests a strong self-compensation effect. This effect is usually associated with the amphoteric character of silicon atoms and their tendency to form complex defects.¹⁶ An alternative mechanism, based on the unintentional incorporation of impurities, was rejected since the samples were grown under the same growth conditions. We have also rejected the DX center as an alternative due to the small increment of the Hall density (only about $0.5 \times 10^{12} \text{ cm}^{-2}$) after intensive illumination at 4 K. However, one should be careful when comparing the Hall density and silicon concentration values, because the Hall density is a function of the concentrations and mobilities of all populated subbands.¹⁷

Qualitatively, the behavior of the Hall density presents similar behavior for the three samples. The samples show a very low variation (less than one order of magnitude) under a high temperature range (150–300 K) and the curves do not

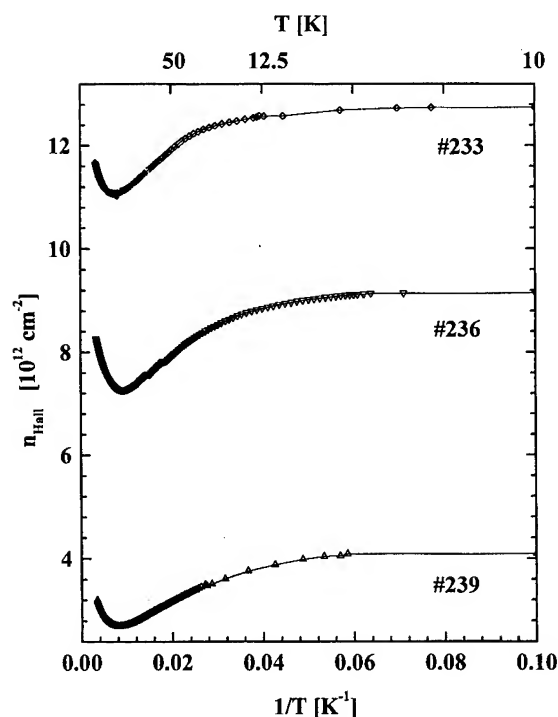


FIG. 1. Temperature dependence of the Hall density for the samples No. 236 (downward-oriented triangle), No. 233 (diamond) and No. 239 (upward-oriented triangle). The Hall effect measurements were obtained in darkness during cooling down of the samples.

fit a typical Arrhenius plot, which could be used to extract a thermal activation energy, eventually related to the DX center.¹⁸ Below 150 K the Hall density increases for decreasing temperature, thus at low temperatures it is higher than at room temperature. These facts suggest that the DX center, if present at all, has different electrical transport properties than those observed in uniformly doped samples. The absence of the freezing out effect observed at low temperature (<150 K) is reported in the literature on delta-doped GaAs samples.¹⁹ In fact, the freezing of electrons in shallow donor levels is rejected, since the Mott transition (about 10^{11} cm⁻²) might have already occurred. Söderström *et al.*²⁰ observed similar behavior of the Hall density in an undoped InSb sample for a layer thickness of less than 5 μ m. In this case, the explanation of this behavior was related to the competition between carriers generated in the regimes of intrinsic and extrinsic conductivities. In our samples the contribution due to intrinsic carriers may be negligible due to the large energy gap. We have proposed that the effect is associated with the redistribution of populations between the subbands in the delta-doped well which does not demand the presence of a DX center.²¹ This assumption is reinforced by the fact that we have not observed changes in our results when the cooling rate was varied within the range 1–30 K/min.²²

B. *n* to *p* type transition under illumination

An unexpected feature of the samples is the change of the sign of the Hall effect when the experiments were done under strong illumination with the infrared LED. This effect is

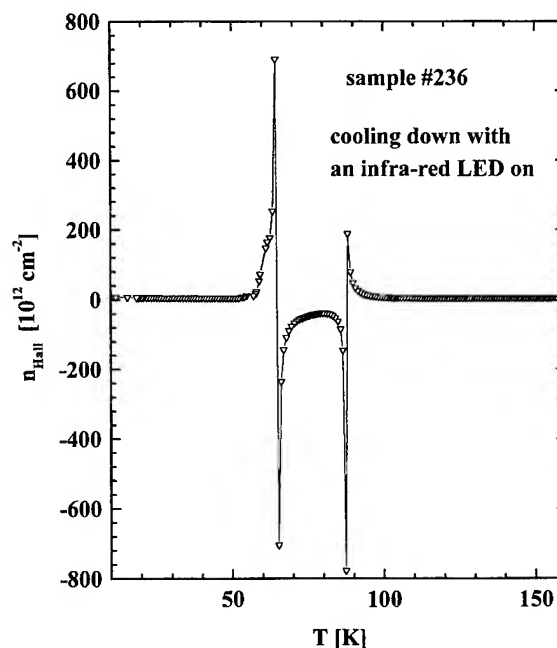


FIG. 2. Temperature dependence of the Hall density under continuous illumination for sample No. 236. The Hall measurement was obtained during cooling down of the sample using an infrared LED of GaAs. We have adopted positive values to the *n*-type characteristic.

illustrated for sample No. 236 in Fig. 2 for a typical cooling down experiment. The positive values correspond to *n*-type characteristics. The sign switch effect was observed in a temperature range of about 90–50 K or, in the case of isothermal illumination, at any temperature below 90 K. An *n*- to *p*-type transition was also observed by Songpongs *et al.*²³ on selectively doped InSb heterostructures grown on GaAs by MBE at about the same temperature range. The authors invoked three conduction channels and a deep defect at the InSb/GaAs interface in order to explain their results, but they did not clearly describe their experimental conditions. Similar transitions were observed previously for single silicon delta-doped GaAs samples.²⁴ The effects are, therefore, intrinsic properties of the delta-doping structure and do not seem to be associated with the DX center.

C. Effect of the cap layer thickness

Because most of the effects observed under permanent illumination are related to the presence of two nearly spatially separated channels,²⁴ we decided to study the electrical properties by varying the distance between the delta-doped layer and the sample surface. The two channels are formed as a 2D electron gas in the delta-doped region (*n*-type channel) and as a hole gas localized just beneath the sample surface (*p*-type channel). We achieved our objective by etching the AlGaAs cap layer, step by step. The eventual effects arising from the nonhomogeneity in different pieces of the same wafer was able to therefore be eliminated. Each etching step was carried out by dipping the sample into a standard etching solution of H₂SO₄:H₂O₂:H₂O (1:8:1000) for 2.5 min. This might correspond to corrosion of about 0.1 μ m of the AlGaAs layer. Figure 3 shows the temporal dependence

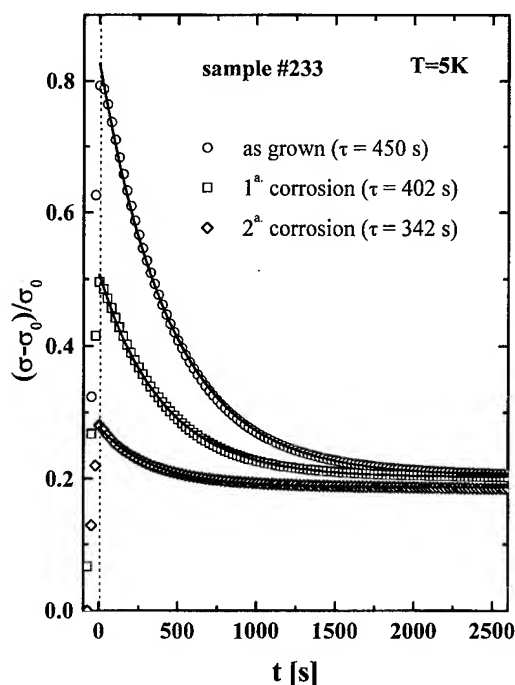


FIG. 3. Relative variation of the Hall conductivity for sample No. 233 under continuous illumination for three different distances of the delta-doped layer from the sample surface. The measurements were obtained using an infrared LED of GaAs operating with 1 mA at 5 K.

of the relative variation of the Hall conductivity ($\Delta(t)/\sigma_0$) for sample No. 233, measured under continuous illumination using the infrared LED, where $\Delta(t) = \sigma(t) - \sigma_0$, $\sigma(t)$ is the conductivity under illumination and σ_0 is the conductivity before illumination. We can observe a reduction of $\Delta(t)$ values when the two conduction channels approach each other. The time relaxation τ of these transients was obtained using the following equation:

$$\Delta(t)/\sigma_0 = \exp(-t/\tau).$$

We chose a weak intensity illumination (1 mA on a LED) in order to avoid nonlinear effects, because otherwise changes in the Hall conductivity would depend intrinsically on the Hall mobility. Immediately after the light was switched on, we observed that the Hall mobility underwent a change of about 15% corresponding to change of about 70% in the Hall density. A clear trend for the relaxation time τ , namely, that it decreases by decreasing the cap layer, suggests that the carriers are strongly localized in the two channels at the limit of faint illumination intensity. The observation of a faint PPC effect corroborates this assumption, and it is indeed related to the spatial separation of charges, rather than with the DX center. The recombination time of charges spatially separated after illumination depends on their separation.²⁵ This kind of persistence effect was also observed in MDS samples.³ Assuming that the persistence is due to electrons released by DX centers, the usual relative concentration of donors acting as DX centers, i.e. $(n_{\text{after}} - n_{\text{dark}})/n_{\text{after}}$, is only about 4% at 5 K, and it decreases still further after each etching step. The values for n_{dark} and

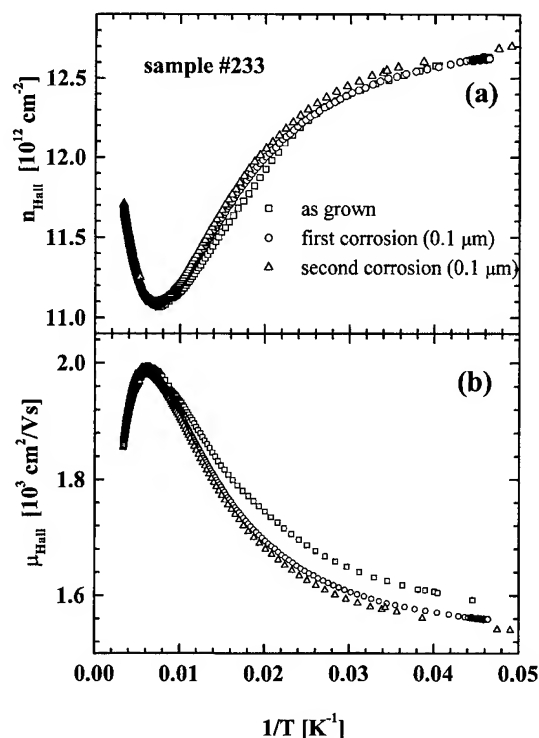


FIG. 4. Temperature dependence of the Hall density (a) and mobility (b) for sample No. 233 in darkness for three different distances of the delta-doped layer from the sample surface.

n_{after} are the Hall densities obtained in darkness and after illumination, respectively. They would be the lowest values up to now reported. These values are about six times lower than those obtained at 77 K by Etienne *et al.*¹ on samples having an electron concentration around 10^{13} cm^{-2} .

Figure 4 shows the temperature dependence of the Hall density and mobility (μ_{Hall}) in darkness for the etched sample. Both the Hall density and the mobility do not change significantly after each etching step. It is well known that the Fermi level is strongly dependent on the defect and impurity state densities formed on the surface.²⁶ Self-consistent calculations solving the Schrödinger and Poisson equations and using the Hall density at low temperature (5 K) as input have shown that the occupation of the subbands is not affected by varying the distance between the spike of delta-doped layer and the sample surface, in agreement with our experimental data (for a more precise procedure one should use the total electron concentration obtained by the Shubnikov-de Haas methods). For the calculations we have assumed an impurity background of 10^{15} cm^{-3} and the Fermi level pinned on the surface at 0.8 eV below conduction band. For any spread of the silicon profile smaller than 100 Å the changes in the computed values are negligible. In sum, the effects of the depletion layer can be ignored, and the surface only plays an important role in the electrical transport properties under illumination.

D. Persistent photoconductivity effect

The PPC effect was observed only as a faint effect as can be seen in Fig. 5 for sample No. 236, and even then only

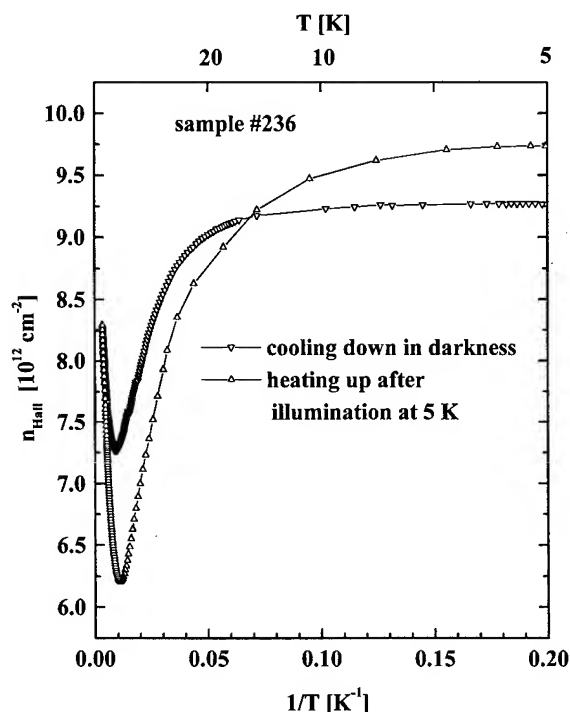


FIG. 5. Temperature dependence of the Hall density for cooling sample No. 236 in darkness (downward-oriented triangle) and after illumination at low temperature heating it using an infrared LED of GaAs (upward-oriented triangle).

for temperatures below 15 K. The downward- and upward-oriented triangles correspond to the Hall density in darkness and after illumination, respectively. The former curve was obtained by cooling down the sample, while the latter was obtained by illuminating the sample at low temperature with high light intensity, then waiting for stabilization of the Hall density, followed by turning off the light source, and finally by heating up the sample. Unexpected is the fact that, for most of the temperature range ($T > 15$ K), the curve “after illumination” lies below the curve obtained in darkness, characterizing a negative PPC effect. These results confront us with three different points, namely (a) the negative PPC effect, for temperatures higher than 15 K, (b) the switch of the negative PPC effect to the positive one at low temperatures, and (c) their relationship to the DX center.

Solomon *et al.*⁸ observed a temperature shift in the threshold of the PPC effect from 120 K to 50 K by replacing the uniformly doped layer by a delta-doped one. This result suggests a lowering of the capture barrier of the DX center. The negative PPC effect was reported by Koenraad *et al.*¹⁹ for delta-doped GaAs samples. Other authors²⁷ observed a negative PPC effect in MDS samples grown by metalorganic vapor phase epitaxy (MOVPE) at a temperature range of 170–300 K. Based on the variation of the energy of the photon (spectral distribution), they correlated the phenomenon to the presence of EL2 centers. EL2-like defects are believed to exist in uniformly doped AlGaAs samples grown by MBE²⁸ as well as in AlGaAs samples using *n-i-n* structures,²⁹ and their concentrations depend on the growth temperature.

Thus, we can argue that the suppression of DX-center features is based on the competition between a relaxed configuration (DX center) and an unrelaxed configuration (EL2-like center) since the two defects show some features in common.³⁰ When low growth temperatures are used, the density of the EL2-like center increases, and its features dominate. The competition could also account for the negative PPC effect. EL2-like defects are invoked to explain an enhancement in the conductivity when silicon delta-doped GaAs samples are illuminated using an infrared LED.³¹ These defects provide a condition for the formation of the two spatially separated channels by means of the transitions: $EL^+ \rightarrow EL^0 \rightarrow EL^{0*}$, where EL^+ , EL^0 and EL^{0*} are the neutral ground, the positively charged and the metastable neutral states, respectively. The photo-Hall density depends on the mobilities of the carriers in the two channels, and because they have opposite *n*- and *p*-type characteristics, the negative persistent effect is possible. In this case, the recombination occurs through tunneling, which does not depend on temperature, thus explaining the maintenance of the negative PPC effect even for temperatures near 250 K.

The results already presented for samples grown at 530 °C suggest that the DX center does not contribute to, or at least does not play a relevant role in, the electron transport properties of silicon delta-doped $Al_{0.30}Ga_{0.70}As$ structures. In order to check this point further we have grown sample No. 231, which was grown at 600 °C (at this temperature we expect a large doping profile beyond 300 Å,³ the ideal delta-doped structure so far). This sample presents high resistivity and *p*-type characteristics in the temperature range from 300 to 50 K. Equivalent enhancement in the resistivity was also observed by other authors in delta-doped structures.³² A possible explanation for the *p*-type characteristics can be based on the assumption that the Mott condition is not reached due either to the high DX center concentration, or the small reduction of the silicon sticking coefficient or the increase of the unintentional impurities. The transport properties should provide a relatively important contribution arising from the background impurity, most probably carbon atoms.

Finally, a point that deserves some consideration regards the growth conditions. The growth interruption just before and after deposition of the silicon atoms was, in an introductory way, considered by Muñoz *et al.*³³ They reported a reduction of the relative silicon concentration acting as the DX center by introducing a 15 s growth interruption. We used a 1 min interruption, so that it might play a role in lowering the concentration of the DX center. Roberts *et al.*³⁴ observed that the DX-center level is strongly dependent on the growth temperature, doping concentration, arsenic overpressure and growth interruption in delta-doped AlGaAs/GaAs quantum well structures. Thus, the delta-doped samples depend critically on the growth parameters.

IV. SUMMARY

Single silicon delta-doped $Al_{0.30}Ga_{0.70}As$ samples were grown by MBE at 530 °C. We observed that in darkness the Hall density does not depend on the cooling rate, at least in

the range of 1–30 K/min. Below 150 K the Hall density increases by decreasing the temperature when in darkness. This was attributed to a redistribution of the population of the delta-doped subbands. The persistent photoconductivity effects, both positive and negative, were associated to the phenomenon of spatial separation of charges, so the pumping mechanisms of electrons into the conduction band and of holes into the valence band are important in defining the electron transport properties of the samples. The absence of DX centers, or at least their low efficiency, was associated to the conduction band being unconnected with the DX-center level. This effect seems to be strongly dependent on the growth parameters. This fact suggests that the DX-center features, usually associated with delta-doped layers, arise from the spread of the silicon profile, which introduces the 3D features of the DX center. The 3D features are stressed by growing the sample at high temperatures due to diffusion effects. The introduction of a delta-doped layer in order to reduce the DX concentration proves to be right, since the growth conditions are adequate to produce an almost ideal delta-doped layer. In summary, it seems that the DX center does not exist in silicon delta-doped AlGaAs samples, which could be explained if we assumed that the DX-center level is not affected by the band bending arising from delta-doped layer, thus making it stay resonant with the conduction band.

ACKNOWLEDGMENTS

The authors would like to thank R. M. Rubinger and K. Krambrock for their valuable discussions on the EL2 center. This work was supported by CAPES, CNPq, FINEP and FAPEMIG.

- ¹B. Etienne and V. Thierry-Mieg, *Appl. Phys. Lett.* **52**, 1237 (1988).
- ²K. Ploog, M. Hauser, and A. Fischer, *Appl. Phys. Lett.* **45**, 233 (1988).
- ³A.-M. Lanzillotto, M. Santos, and M. Shayegan, *Appl. Phys. Lett.* **55**, 1445 (1989).
- ⁴E. F. Schubert, C. W. Tu, R. F. Kopf, J. M. Kuo, and L. M. Lunardi, *Appl. Phys. Lett.* **54**, 2592 (1989).
- ⁵G. M. Yang, S. G. Park, K. S. Seo, and B. D. Choe, *Appl. Phys. Lett.* **60**, 2380 (1992).
- ⁶D. M. Ritchie, A. Di Paola, M. Tromby, M. Dellagiovanna, M. Di Egidio, and F. Vidimari, *J. Cryst. Growth* **145**, 447 (1994).
- ⁷H. Mejri, A. Selmi, H. Maaref, and J. S. Bourgoin, *J. Appl. Phys.* **69**, 4060 (1991).
- ⁸G. S. Solomon, G. Roos, and J. S. Harris, Jr., *J. Cryst. Growth* **127**, 737 (1993).
- ⁹G. M. Yang, K. S. Seo, and B.-D. Choe, *J. Appl. Phys.* **77**, 2046 (1995).
- ¹⁰J. Sánchez-Dehesa, D. Lavielle, E. Ranz, B. Goutiers, J. C. Portal, E. Barbier, A. Y. Cho, and D. L. Sivco, *Semicond. Sci. Technol.* **6**, 445 (1991).
- ¹¹S. Arscott, M. Missous, and I. Dobaczewski, *Semicond. Sci. Technol.* **7**, 620 (1992).
- ¹²J. C. Bourgoin, S. L. Feng, and H. J. von Bardeleben, *Phys. Rev. B* **40**, 7663 (1989).
- ¹³D. J. Chadi and K. J. Chang, *Phys. Rev. B* **39**, 10 063 (1989).
- ¹⁴E. F. Schubert, J. M. Kuo, R. F. Kopf, A. S. Jordan, H. S. Luftman, and L. C. Hopkins, *Phys. Rev. B* **42**, 1364 (1990).
- ¹⁵G. Brunthaler, G. Stöger, A. Aumayr, and K. Köhler, *Appl. Phys. Lett.* **62**, 1635 (1993).
- ¹⁶A. G. de Oliveira, G. A. M. Safar, and J. F. Sampaio, *Braz. J. Phys.* **24**, 363 (1994).
- ¹⁷I. A. Panaev, S. A. Studenikin, D. I. Lubyshev, and V. P. Migal, *Semicond. Sci. Technol.* **8**, 1822 (1993).
- ¹⁸N. Chand, T. Henderson, J. Klen, W. T. Massselink, R. Fischer, Y. C. Chang, and H. Morkoç, *Phys. Rev. B* **30**, 4481 (1984).
- ¹⁹P. M. Koenraad, F. A. P. Blom, C. J. G. M. Langerak, M. R. Leys, J. A. A. J. Perenboom, J. Singleton, S. J. R. M. Spermon, W. C. van der Vleuten, A. P. J. Voncken, and J. H. Wolter, *Semicond. Sci. Technol.* **5**, 861 (1990).
- ²⁰J. R. Söderström, M. M. Cumming, J.-Y. Yao, and T. G. Andersson, *Semicond. Sci. Technol.* **7**, 337 (1992).
- ²¹J. A. Corrêa F., A. G. de Oliveira, M. I. N. da Silva, M. V. B. Moreira, G. M. Ribeiro, and H. Chacham, *Solid State Commun.* **98**, 1063 (1996).
- ²²A. Y. Du, M. F. Li, T. C. Chong, and S. J. Chua, *Appl. Phys. Lett.* **66**, 1391 (1995).
- ²³P. Songpongs, T. G. Andersson, M. J. Ekenstedt, J. R. Söderström, and M. M. Cumming, *Appl. Phys. Lett.* **65**, 1433 (1994).
- ²⁴A. G. de Oliveira, G. M. Ribeiro, D. A. W. Soares, J. A. Corrêa F., M. I. N. da Silva, and H. Chacham, *J. Appl. Phys.* **78**, 2659 (1995).
- ²⁵H. J. Queisser and D. E. Theodorou, *Phys. Rev. Lett.* **43**, 401 (1979).
- ²⁶T. M. Hsu, W. C. Lee, J. R. Wu, and J.-I. Chyi, *Phys. Rev. B* **51**, 17 215 (1995).
- ²⁷H. Pettersson, H. G. Grimmeiss, A. L. Powell, C. C. Button, J. S. Roberts, and P. I. Rockett, *J. Appl. Phys.* **74**, 5596 (1993).
- ²⁸K. Yamanaka, S. Naritsuka, K. Kanamoto, M. Mihara, and M. Ishii, *J. Appl. Phys.* **61**, 5062 (1987).
- ²⁹A. K. Verma, J. Tu, J. S. Smith, H. Fujioka, and E. R. Weber, *Appl. Phys. Lett.* **68**, 699 (1996).
- ³⁰P. Hautojärvi, *J. Phys. (Paris), Colloq.* **5**, 1 (1995).
- ³¹R. M. Rubinger, A. G. de Oliveira, M. I. N. da Silva, J. A. Corrêa F., M. V. B. Moreira, G. M. Ribeiro, and H. Chacham, 23rd International Conference on the Physics of Semiconductors, Vol. VI.C.15, 2797 (1996).
- ³²M. C. Holland, E. Skuras, J. H. Davies, I. A. Larkin, A. R. Long, and C. R. Stanley, *J. Cryst. Growth* **150**, 1215 (1995).
- ³³E. Muñoz, E. Calleja, I. Izpura, F. García, A. L. Romero, J. L. Sánchez-Rojas, A. L. Powell, and J. Castagné, *J. Appl. Phys.* **73**, 4988 (1993).
- ³⁴J. M. Roberts, J. J. Harris, and C. Roberts, *Semicond. Sci. Technol.* **11**, 458 (1996).

X-ray photoelectron spectroscopy study of GaAs(110) cleaved in alcoholic sulfide solutions

Vasily N. Bessolov, Elena V. Konenkova, and Mikhail V. Lebedev^{a)}

A. F. Ioffe Physico-Technical Institute, Russian Academy of Sciences, Politekhnikeskaya 26, St. Petersburg 194021, Russia

(Received 23 January 1997; accepted 4 April 1997)

X-ray photoelectron spectroscopy has been used to study chemical bonds which appeared on the (110) surface of GaAs after cleavage of the crystal in solutions of sodium sulfide in different alcohols. It has been shown that after such treatment the sulfur atoms are almost exclusively bonded only to gallium atoms. This is conditioned by the electrostatic interaction of sulfur ions in the solution with the surface ions of the crystal lattice of the semiconductor. © 1997 American Vacuum Society. [S0734-211X(97)00504-0]

I. INTRODUCTION

The surface of III–V semiconductors is covered as usual by a relatively thick native oxide layer which results in a large density of surface states in the middle of the gap and a high surface recombination velocity which is detrimental to the performance of microelectronic and optoelectronic devices.

Treatment of the semiconductors in sulfide solutions results in an effective surface passivation. In particular, such treatment proved to be effective in removal of the native oxide layer and in reducing the surface recombination velocity and surface states density. By sulfide passivation the performance of both transistors^{1,2} and lasers^{3–8} could be improved dramatically.

The modification of the semiconductor surface in the course of solution sulfidizing is a result of the chemical interaction of sulfur ions with the surface. It was shown^{9,10} that the efficiency of the GaAs(110) surface passivation could be considerably increased through the use of solvents in sulfide solutions of various alcohols with low dielectric constant value (e.g., isopropanol, *tert*-butanol, etc.). Moreover it appeared that the efficiency of GaAs(100) electronic passivation is enhanced when sulfidizing processes with higher rate constants are used.¹¹

Unlike (100) surfaces there are very few articles concerned with sulfur treatment of the (110) surface of III–V semiconductors. By Auger electron and x-ray photoelectron spectroscopies it was shown, first, that the treatment in ammonium sulfide aqueous solution of epitaxially grown AlGaAs(110) layer with native oxide results in the decrease of the surface oxides content and the appearance of sulfides (mainly aluminum sulfides).¹² Second, it was found¹³ that adsorption of H₂S on cleaved in vacuum (110) surfaces of InP, GaP, and GaAs results in the formation of sulfides of V group atoms and, what is more, for H₂S adsorption on GaAs(110) surface a reduced processing temperature ($T = 200$ K) is required. On the other hand, the sulfur treatment of mirrors of semiconductor lasers [i.e., just the (110) surface] in both sodium sulfide and ammonium sulfide aqueous

and alcoholic solutions markedly increases their catastrophic optical damage level^{3,6–8} due to the reduction of the surface recombination velocity on the mirrors.

In this article we investigated by x-ray photoelectron spectroscopy the chemical bonds on the (110) surface of GaAs cleaved in alcoholic sulfide solutions.

II. EXPERIMENT

The investigations were performed on *n*-GaAs:Sn ($n = 5 \times 10^{17}$ cm⁻³) 2-mm-thick crystals. The crystal was put into sulfide solution and cleaved there. The cleaved (110) surface was exposed in the solution for 1 min and dried in air.

The sulfide treatment was carried out under illumination with a 200 W incandescent lamp at room temperature. The solutions used were saturated solutions of sodium sulfide (Na₂S: x H₂O with 60%–62% of Na₂S) in different alcohols with low dielectric constant value ϵ : in isopropanol [*i*-C₃H₇OH ($\epsilon = 18.3$)], *tert*-butanol [*t*-C₄H₉OH ($\epsilon = 11$)], and 2-methylbutanol-2 [*t*-C₅H₁₁OH ($\epsilon = 5.85$)]. The estimated sulfur contents in different solutions are listed in Table I.

To investigate the GaAs surface composition x-ray photoelectron spectroscopy was used. Photoelectron spectra were recorded with a Perkin–Elmer PHI 5400 spectrometer using a Mg K_{α} ($h\nu = 1253.6$ eV) x-ray source operating at 300 W. The samples were put into the vacuum chamber not later than 15 min after finishing the sulfidizing process. Data analysis was carried out using Perkin–Elmer software.

III. RESULTS

The survey photoelectron spectra of GaAs(110) treated surfaces contained Ga, As, C, O, Na, and S peaks. Sulfidized surfaces were Ga-rich. The highest Ga/As ratio (1.36) was observed for the surface treated in *i*-C₃H₇OH-based solution while the lowest one (1.14) was for the surface treated in *t*-C₅H₁₁OH-based solution. Therefore the surface obtained was more stoichiometric in solution with the lowest dielectric constant.

Figures 1 and 2 show high resolution spectra of Ga 3*d* and As 3*d* peaks for sulfidized surfaces. Relative intensities of different peaks and sulfur coverage for each GaAs(110)

^{a)}Electronic mail: nsab@les.ioffe.rssi.ru

TABLE I. The composition of the solution.

Solution	Sulfur content in the solution ($\text{g} \times \text{cm}^{-3}$)
$\text{Na}_2\text{S} + i\text{-C}_3\text{H}_7\text{OH}$	1.6×10^{-6}
$\text{Na}_2\text{S} + t\text{-C}_4\text{H}_9\text{OH}$	2.4×10^{-6}
$\text{Na}_2\text{S} + t\text{-C}_5\text{H}_{11}\text{OH}$	$10^{-8} \times 10^{-9}$

surface are listed in Table II. It is seen that in spite of a great difference in sulfur content in the solutions (several orders of magnitude) the sulfur coverages for all treated surfaces do not differ so much.

The Ga 3d spectra of GaAs(110) surfaces treated in all sulfide solutions can be fitted very well with three components (Fig. 1), namely, the first is a Ga-As bond with binding energy 19.6 eV, the second shifted by 0.73 eV to higher binding energies, and the third shifted by 1.45 eV. Taking into account the well-known chemical shifts observed in the photoelectron spectra of oxidized and sulfur passivated GaAs surfaces^{14,15} and the presence in the spectra of the S 2s peak as well (Fig. 3 shows typical S 2s peak which had the same shape but different intensities for all samples considered in this study), the second component could be associated with Ga sulfides (the Ga-S bonds) and the third one could be associated with the Ga oxides (the Ga-O bonds). The Ga-S/Ga-O ratio increases with the decrease of the dielectric constant of the solvent from 0.87 for a surface treated in

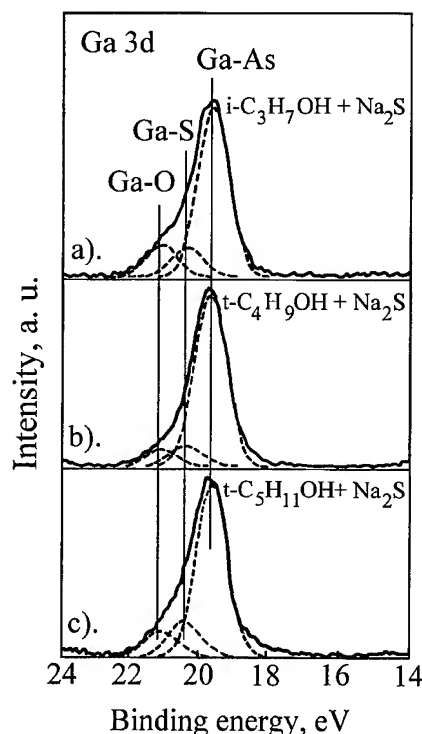


FIG. 1. Ga 3d x-ray photoelectron spectra of GaAs(110) cleaved and exposed at room temperature for 1 min in different saturated sodium sulfide solutions. (a) In the isopropanol-based solution, (b) in the *tert*-butanol-based solution, (c) in the 2-methylbutanol-2-based solution.

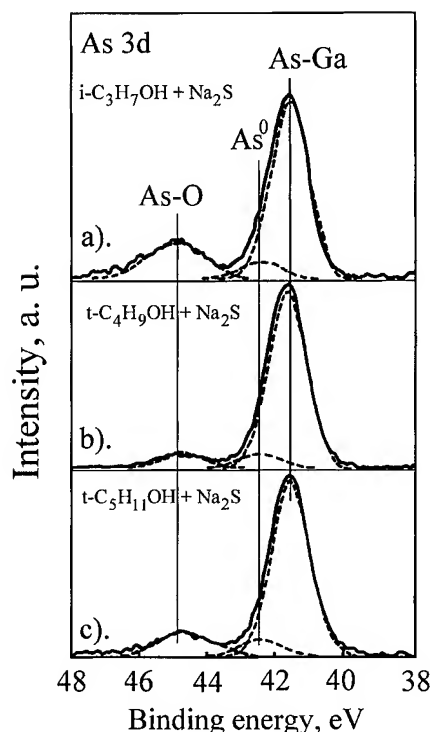


FIG. 2. As 3d x-ray photoelectron spectra of GaAs(110) cleaved and exposed at room temperature for 1 min in different saturated sodium sulfide solutions. (a) In the isopropanol-based solution, (b) in the *tert*-butanol-based solution, (c) in the 2-methylbutanol-2-based solution.

i-C₃H₇OH-based solution up to 1.31 for a surface treated in *t*-C₅H₁₁OH-based solution.

The As 3d spectra of GaAs(110) surfaces treated in all sulfide solutions can be fitted very well with three components (Fig. 2), namely, the first is an As-Ga bond with binding energy 41.6 eV, the second shifted by 0.85 eV to higher binding energies, and the third shifted by 3.20 eV. Taking into account the well-known chemical shifts observed in the photoelectron spectra of oxidized and sulfur passivated GaAs surfaces,^{14,15} the second component could be associated with elemental As (As⁰) and the third one could be associated with the As oxides (the As-O bonds). It should be noted that the elemental As⁰ content is the same for all (110) surfaces considered in this study.

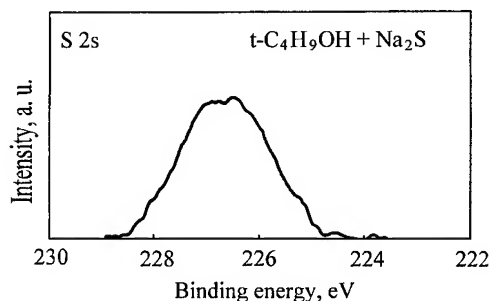


FIG. 3. Typical S 2s x-ray photoelectron spectra of GaAs(110) cleaved and exposed at room temperature for 1 min in saturated sodium sulfide *tert*-butanol-based solution.

TABLE II. Relative intensities of different peaks and sulfur coverage for GaAs(110) cleaved in different solutions.

Solution	S 2s	As-O	As-S	As ⁰	Ga-O	Ga-S	Sulfur coverage (MLs)
	Ga 3d+As 3d	Ga 3d+As 3d	Ga 3d+As 3d	Ga 3d+As 3d	Ga 3d+As 3d	Ga 3d+As 3d	
Na ₂ S + <i>i</i> -C ₃ H ₇ OH	0.07	0.09	...	0.03	0.08	0.07	1.04
Na ₂ S + <i>t</i> -C ₄ H ₉ OH	0.05	0.04	...	0.03	0.05	0.05	0.74
Na ₂ S + <i>i</i> -C ₅ H ₁₁ OH	0.09	0.07	...	0.03	0.06	0.09	1.34

IV. DISCUSSION

First of all it should be stressed that on the GaAs(110) surface cleaved in alcoholic sulfide solution the sulfur is exclusively bonded only to Ga atoms and not to As atoms at all, while for oxidized GaAs(100) surface treated in alcoholic solutions of sodium sulfide the sulfur is bonded mainly to As atoms.^{9,10} Such different behavior of sulfur on these surfaces could be explained by taking into account the electrostatic interaction between sulfur ions and the semiconductor surface which plays an important role in the solution sulfidizing process.^{9,10,16} Since, on the freshly cleaved surface there is no band bending at least immediately after the cleavage, the exclusive formation of the Ga-S bonds on GaAs(110) surface cleaved in the solution apparently results from the attraction of the negative sulfur ion to the positive Ga ion appearing on the cleaved surface and, simultaneously, from the sulfur ions, being repelled from the negative As ion. The primary formation of the As-S bonds on the oxidized GaAs(100) surface, in its turn, apparently results from the interaction of the sulfur ion with the surface dipole layer related with band bending and/or foreign adatoms. In that case, according to the hard and soft acids and bases principle,¹⁷ the sulfur ion which is the soft base should try to form the covalent bonds with the softest atom on the surface that is with As¹⁸ in spite of the higher thermodynamic stability of the Ga-S bonds compared with the As-S bonds.

The effective rate constant of the reaction of sulfur coat formation, which was estimated as the sulfur coverage formed on the GaAs(110) surface after exposure in the solution for 1 min (Table II) divided by sulfur content in the solution (Table I), is increased with reciprocal dielectric constant of the solvent (Fig. 4) like in the case of GaAs(100) sulfur treatment.^{9,10} Such dependence also gives evidence in favor of the importance of the electrostatic interaction in the process of sulfur ions adsorption on the (110) surface of GaAs.

The existence of oxides and elemental As⁰ on the GaAs(110) surface most likely results from the surface oxidation during its transfer to vacuum chamber. Indeed, as it was shown^{14,19} exposure to air of a sulfidized surface (in particular treated in sodium sulfide solutions) results in rapid formation of the Ga and As oxides. On the other hand, the treatment of the GaAs(100) natively oxidized surface in alcoholic solutions of sodium sulfide results in the reduction of the total amount of oxides and has not effect on the elemental As⁰ content.⁹ The latest fact testifies that the sulfur treatment process should not introduce additional oxides and elemental As⁰.

It should be mentioned that the other techniques for sulfidizing of different GaAs surfaces results in primary formation of the As-S bonds as a rule (the Ga-S bonds content is usually less) while a surface covered exclusively by Ga-S bonds could be obtained only after vacuum annealing of the sulfur terminated surface due to thermoactive conversion of the As-S bonds to the Ga-S bonds.²⁰ Nevertheless, the exclusive formation of the Ga-S bonds resulted from the exposure of GaAs(100) to gas phase atomic sulfur at 60–70 °C.²¹

V. SUMMARY

In this article chemical bonds appearing on the (110) surface of GaAs after cleavage of the crystal in a solution of sodium sulfide in different alcohols (isopropanol, *tert*-butanol, 2-methylbutanol-2) were studied. It was shown that after such treatment the sulfur atoms are almost exclusively bonded only to gallium atoms. This is conditioned by the

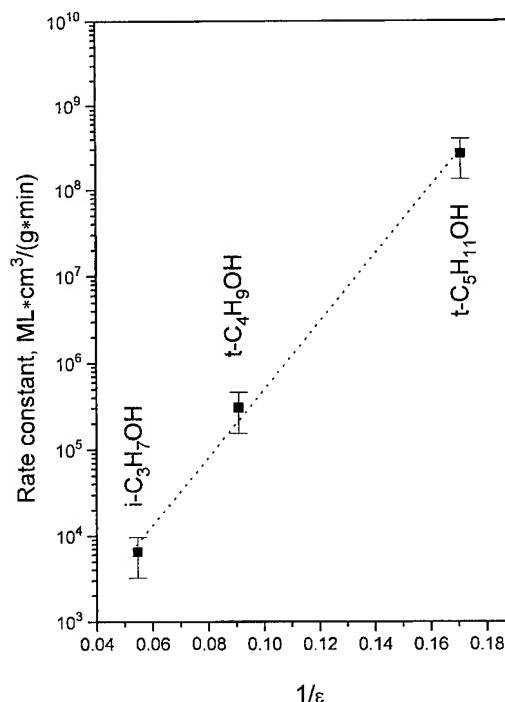


FIG. 4. The estimated values of the effective rate constant of the reaction of sulfur coat formation on GaAs(110) surface cleaved and exposed at room temperature for 1 min in different sulfide solutions vs the reciprocal dielectric constant of the solvent used.

electrostatic interaction of sulfur ions in the solution with the surface ions of the crystal lattice of the semiconductor.

ACKNOWLEDGMENTS

The research described in this article was made possible in part by Grant No. 95-03-09330 from the Russian Foundation for Basic Research, and by a grant from Volkswagen-Stiftung.

- ¹R. N. Nottenburg, C. J. Sandroff, D. A. Humphrey, T. H. Hollenbeck, and R. Bhat, *Appl. Phys. Lett.* **52**, 218 (1988).
- ²J.-L. Lee, D. Kim, S. J. Maeng, H. H. Park, J. Y. Kang, and Y. T. Lee, *J. Appl. Phys.* **73**, 3539 (1993).
- ³S. Kamiyama, Y. Mori, T. Takahashi, and K. Ohnaka, *Appl. Phys. Lett.* **58**, 2595 (1991).
- ⁴L. F. DeChiaro and C. J. Standroff, *IEEE Trans. Electron Devices* **39**, 561 (1992).
- ⁵W. S. Hobson, U. Mohideen, S. J. Pearton, R. E. Slusher, and F. Ren, *Electron. Lett.* **29**, 2199 (1993).
- ⁶A. J. Howard, C. I. H. Ashby, J. A. Lott, R. P. Schneider, and R. F. Corless, *J. Vac. Sci. Technol. A* **12**, 1063 (1994).
- ⁷V. N. Bessolov, M. V. Lebedev, Y. M. Shernyakov, B. V. Tsarenkov, J. S. Yu, J. H. Lee, G. T. Pak, and T. I. Kim, *Tech. Phys. Lett.* **21**, 561 (1995); V. N. Bessolov, M. V. Lebedev, Y. M. Shernyakov, and B. V. Tsarenkov, *Mater. Sci. Eng. B* **44**, 380 (1997).
- ⁸G. Beister, J. Maegle, D. Gutsche, G. Erbert, J. Sebastian, K. Vogel, M. Weyers, J. Wurfl, and O. P. Daga, *Appl. Phys. Lett.* **68**, 2467 (1996).
- ⁹V. N. Bessolov, E. V. Konenkova, and M. V. Lebedev, *J. Vac. Sci. Technol. B* **14**, 2761 (1996).
- ¹⁰V. N. Bessolov, E. V. Konenkova, and M. V. Lebedev, *Mater. Sci. Eng. B* **44**, 376 (1997).
- ¹¹V. N. Bessolov, E. V. Konenkova, and M. V. Lebedev, *Tech. Phys. Lett.* **22**, 749 (1996).
- ¹²H. Ohno, H. Kawanishi, Y. Akagi, Y. Nakajima, and T. Hijikata, *Jpn. J. Appl. Phys.* **29**, 2473 (1990).
- ¹³E. Dudzik, C. Muller, I. T. McGovern, D. R. Lloyd, A. Patchett, D. R. T. Zahn, T. Johal, and R. McGrath, *Surf. Sci.* **344**, 1 (1995).
- ¹⁴T. Ishikawa and H. Ikoma, *Jpn. J. Appl. Phys.* **31**, 3981 (1992).
- ¹⁵K. Sato, M. Sakata, and H. Ikoma, *Jpn. J. Appl. Phys.* **32**, 3354 (1993).
- ¹⁶V. N. Bessolov, A. F. Ivankov, and M. V. Lebedev, *J. Vac. Sci. Technol. B* **13**, 1018 (1995).
- ¹⁷R. G. Pearson, *J. Am. Chem. Soc.* **85**, 3533 (1963).
- ¹⁸P. Piquini, A. Fazzio, and A. Dal Pino, Jr., *Surf. Sci.* **313**, 41 (1994).
- ¹⁹M. Oshima, T. Scimeca, Y. Watanabe, H. Oigawa, and Y. Nannichi, *Jpn. J. Appl. Phys.* **32**, 518 (1993).
- ²⁰H. Sugahara, M. Oshima, R. Klauser, H. Oigawa, and Y. Nannichi, *Surf. Sci.* **242**, 335 (1991).
- ²¹G. Y. Gu, E. A. Ogryzlo, P. C. Wong, M. Y. Zhou, and K. A. R. Mitchell, *J. Appl. Phys.* **72**, 762 (1992).

Effects of water vapor and chlorine on the epitaxial growth of $\text{Si}_{1-x}\text{Ge}_x$ films by chemical vapor deposition: Thermodynamic analysis

I.-M. Lee, A. Jansons, and C. G. Takoudis^{a)}

School of Chemical Engineering, Purdue University, West Lafayette, Indiana 47907-1283

(Received 6 December 1996; accepted 25 April 1997)

The effects of water vapor and chlorine on the $\text{Si}_{1-x}\text{Ge}_x$ epitaxial growth are studied through thermodynamic analyses. Since Ge does not form chemical species with oxygen at an appreciable level, the water vapor and oxygen background level requirements for $\text{Si}_{1-x}\text{Ge}_x$ epitaxial growth are mostly similar to the ones for Si growth. At large excesses of Ge gas phase sources and low system water vapor level, however, the critical temperature above which oxide-free epitaxial growth occurs decreases with increasing system pressure. Analyses of the results obtained suggest that there are two mechanisms for the removal of $\text{SiO}_{2(s)}$ from the surface: one is by forming volatile SiO , and the other is by forming gas phase H_2O . The former one appears to dominate for most conditions studied; the latter one appears to dominate when the water vapor level of the system is very low and/or small amounts of the Si gaseous source are present. The boundary between etching and deposition of $\text{Si}_{1-x}\text{Ge}_x$ films is examined for all Ge solid phase compositions. The etch/deposition boundaries for pure Si and pure Ge show opposite trends with increasing chlorine concentrations, because of the different affinities of Si and Ge to form chlorinated species. The chlorine concentration needed to reach etching increases with increasing system temperature for Si, whereas it decreases with increasing temperature for Ge. It is found that the formation of SiCl_4 is favored over GeCl_2 at low temperatures, while the opposite happens at higher temperatures. The etch/deposition boundary for $\text{Si}_{1-x}\text{Ge}_x$ thin films is found to form a saddlelike contour because of the interaction of the different behaviors of Si and Ge. Overall, calculated results are seen to be in good agreement with available experimental results. © 1997 American Vacuum Society. [S0734-211X(97)09804-1]

I. INTRODUCTION

Epitaxial growth of strained $\text{Si}_{1-x}\text{Ge}_x$ films is usually carried out at a lower temperature than that of Si films in order to minimize the relaxation of $\text{Si}_{1-x}\text{Ge}_x$ films due to lattice mismatch.¹⁻³ Furthermore, high performance integrated circuits require smaller geometry in lateral as well as vertical dimensions and abrupt transitions in dopant concentration between adjacent single crystalline layers. These requirements exclude conventional high-temperature ($>1000^\circ\text{C}$) epitaxy and low-temperature alternatives are thus called for.

One of the most important aspects of low-temperature epitaxy is the background water vapor and oxygen levels in the growth environment. As demonstrated by Meyerson with an ultrahigh vacuum reactor, the results from Ghidini and Smith^{4,5} can be used to predict an operating range for commercially available epitaxial reactor systems where good crystalline quality material can be grown.^{6,7} For epitaxial $\text{Si}_{1-x}\text{Ge}_x$ growth, where a source of germanium is added to the system, the allowed background water vapor levels do not appear to have been examined systematically. This is one of the thrusts of this study.

Selective epitaxial growth (SEG) is a process which allows the epitaxial film to be grown only in selected regions on a patterned wafer. SEG evolved from full wafer epitaxy and hence the growth conditions are quite similar to those of

full wafer epitaxy with one notable exception, the amount of chlorine in the system. The role of chlorine is to etch possible nuclei of Si and/or Ge atoms on the oxide (or other patterning thin film) surface. However, if excess of chlorine is added, the etching reaction becomes dominant and etching of the Si substrate can thus take place. Therefore, there is a boundary between the regions of deposition and etch.⁸ In the case of $\text{Si}_{1-x}\text{Ge}_x$ -based devices, SEG provides much more flexibility in the design of novel high-speed device structures, limits defect propagation, and reduces misfit and threading dislocation densities.⁹ In this study, the amount of chlorine that defines the etch/deposition boundary with respect to various processing environments is investigated.

As in Ref. 10, the $\text{Si}_{1-x}\text{Ge}_x$ film composition is obtained as a function of reactor temperature, pressure, and reactant composition. Such results provide useful information for reactor design as well as important guidelines for experimental studies. Understanding the effects of oxygen and chlorine on the epitaxial growth of $\text{Si}_{1-x}\text{Ge}_x$ can substantially help with the elucidation of relationships between thin film properties and processing environments in these systems.

II. METHODOLOGY

In a chemical vapor deposition (CVD) process, there are several constraints that must be satisfied. They are total mass conservation, conservation of the initial amounts of atomic

^{a)}Present address: Dept. of Chemical Engineering, University of Illinois at Chicago, Chicago, IL 60607.

TABLE I. Ranges of variables for the water vapor effects on the $\text{Si}_{1-x}\text{Ge}_x$ epitaxial growth.

Variable	Range	Notes
Ge/(Ge+Si)	25%–100%	Each simulation contained a total of 5.5% of the Si/Ge mixture
H_2O concentration	0.002–1.0 ppm	No calculation for 0 ppm since no oxides could occur
Pressure	0.1–10.0 Torr	
Temperature	400–900 °C	

species, constant temperature, and constant total pressure. To carry out a thermodynamic equilibrium calculation, two methods can be used:

- (1) by considering a set of chemical reactions and using the equilibrium constants of these reactions,¹¹ or
- (2) by considering a number of chemical species and minimizing the free energy of the system.¹²

The results from both should be identical. In this study, the second approach is used. Depending on reaction conditions, a system can be either near or away from thermodynamic equilibrium. If the system is close to equilibrium, calculated thermodynamic results should be in satisfactory agreement with corresponding experimental data.

In this study, a computer program originating from NASA was used.¹³ The thermodynamic data of the species involved were fitted to a polynomial format as a function of temperature as required by the computer program. The data needed were heat capacity, entropy, and enthalpy. A list of the species included in this study can be found in Ref. 10. The reliability of these data was tested in studies including subsystems such as Si/H/Cl and Ge/H/Cl. The basic methodology included all conceivable species in the system of interest as well as the initial (feed) conditions; then the system free energy was minimized at the conditions of interest. Results from these analyses were compared with available experimental data,¹⁰ and the agreement between the two was satisfactory.

The parameters studied for oxygen impurity effects on the $\text{Si}_{1-x}\text{Ge}_x$ growth are reported in Table I. These parameters were chosen to replicate conditions used in the low-pressure chemical vapor deposition (LPCVD) of $\text{Si}_{1-x}\text{Ge}_x$ in Ref. 14. The gas species included H_2 , SiH_2Cl_2 , GeH_4 , H_2O , and HCl. In each calculation, the Si/Ge gas mixture was 5.5% of the total input. One specific comment is in order at this point. The computer program only considers the relative amounts of elements present. Therefore, equivalent results would be obtained given equivalent element ratios regardless of the different forms of the species which consist of these elements.

At a system pressure, water vapor level in the reactor, chlorine level, and Si/Ge ratio, the analysis was performed by increasing the system temperature starting from 400 °C in increments of 50 °C until the disappearance of all oxides from the solid phase. Then that particular temperature was

TABLE II. Ranges of variables for the chlorine effect on the $\text{Si}_{1-x}\text{Ge}_x$ growth/etching.

Variable	Range	Notes
Ge/(Ge+Si)	0%–100%	A fixed ratio of (Si+Ge):H=1:50 throughout the analysis
Pressure	0.1–10.0 Torr	
Temperature	500–900 °C	
Cl/H ratio	0.06–0.1	

the critical one above which no solid oxides form.

The effect of chlorine on $\text{Si}_{1-x}\text{Ge}_x$ growth was investigated with the parameters listed in Table II. Again, the parameters were chosen to replicate LPCVD conditions reported in Ref. 14. The ratio of (Si+Ge)/H was kept at 1:50 for all analyses so that the total amount of (Si+Ge) was constant. At a reactor pressure, temperature, and Si/Ge ratio, analyses were performed by increasing the chlorine level until the disappearance of Si and Ge from the solid phase. Then that particular chlorine level was assigned as the critical chlorine level at which no deposition of Si or Ge takes place.

III. RESULTS AND DISCUSSION

A. Effect of water vapor

As indicated in Table I, this study investigates initial Ge contents from 25% to 100%. These results combined with the ones in Ref. 15(a), where conditions for oxide-free Si growth have been studied, cover the entire Ge content range from 0% to 100%.

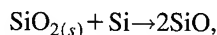
Except for the very rich in Ge mixtures, all calculated results show that as the system pressure and/or the water concentration is increased, the temperature required to avoid solid oxide formation increases (Table III). This trend is in agreement with results reported by Ghidini and Smith^{4,5} and Friedrich *et al.*⁷ as well as by Gaynor^{15(a)} from calculations under similar conditions, in the absence of germanium. A plot of system temperature versus mole fraction of oxygen containing species shows two regions (Fig. 1). Each of these regions corresponds to a temperature range in which one oxygen containing species is preferred. Primarily two chemical species dominate the oxygen containing species, $\text{SiO}_{2(s)}$ at lower temperatures and $\text{SiO}_{(g)}$ at higher temperatures. The trend for all conditions below the 99.5% germanium level is similar to the one shown in Fig. 1 with the exception that the boundary between the $\text{SiO}_{(g)}$ and $\text{SiO}_{2(s)}$ regions is shifted to the right for higher initial concentrations of H_2O and higher system pressure (and vice versa, shifted to the left for lower concentrations of H_2O and lower system pressure).

A different trend is observed in the near absence of silicon. As shown in Table IV, the temperature required to maintain a $\text{SiO}_{2(s)}$ -free surface shows two different trends. For water vapor levels of 0.1 and 1.0 ppm, the required temperature increases with increasing pressure which is in accordance with the trend reported in Table III. However, at very low water vapor level (0.002 ppm), the critical temperature decreases with increasing pressure.

TABLE III. The highest temperature at which solid phase oxides are present for Ge/(Ge+Si) initial ratios between 0% and 99%.

H ₂ O concentration (ppm)	Pressure (Torr)		
	P = 0.1	P = 1.0	P = 10
0.002	600 °C	650 °C	700 °C
0.100	650 °C	700 °C	750 °C
1.000	700 °C	750 °C	800 °C

An explanation based on the calculated composition of major oxygen-containing species is presented next. The removal of SiO_{2(s)} from the substrate surface may be achieved through two mechanisms. In one, SiO_{2(s)} is removed by forming SiO. the following reaction can account for the formation of SiO:



where the silicon in this reaction could come from the solid or gas phase.^{4,5,15(b)} This is denoted as mechanism A in Table V. However, a different scenario may take place when the water vapor content is very low and/or a low level of a Si source is available. SiO_{2(s)} may then leave the surface by reacting with hydrogen so that oxygen atoms are carried away from the substrate surface in the form of H₂O (mechanism B in Table V). This is corroborated by equilibrium composition calculations which show that the formation of SiO is favored with decreasing pressure (Fig. 1), while the formation of H₂O is favored with increasing pressure at the

TABLE IV. The highest temperature at which solid SiO₂ is present for Ge/(Ge+Si) initial ratio of 99.5%.

H ₂ O concentration (ppm)	Pressure (Torr)		
	P = 0.1	P = 1.0	P = 10
0.002	550 °C	450 °C	400 °C
0.100	700 °C	750 °C	750 °C
1.000	750 °C	800 °C	900 °C

same temperature (Fig. 2). Therefore, the temperature required to maintain an oxide-free surface increases with increasing pressure for mechanism A, and vice versa for mechanism B.

By looking at the concentration changes of H₂O, SiO, and SiO_{2(s)}, the region in which one mechanism dominates over the other can be readily seen. For mechanism A in which SiO_{2(s)} is removed by formation of SiO, it is clearly seen that the decrease in SiO_{2(s)} concentration with respect to increasing temperature is accompanied by an increase in SiO concentration (Fig. 1). For mechanism B in which SiO_{2(s)} is removed by formation of H₂O, the decrease in the SiO_{2(s)} concentration is accompanied by an increase in H₂O concentration (Fig. 2). When the temperature is further increased, the H₂O concentration decreases and SiO concentration increases because conditions strongly favor SiO formation. This finding may reconcile, in part, the differences reported by Greve and Racanelli¹⁶ and Meyerson *et al.*⁹ Both groups used similar ultrahigh vacuum CVD systems but one group reported that *in situ* cleaning by a hydrogen bake was needed for quality film deposition,¹⁶ while the other group reported that no *in situ* cleaning by a hydrogen bake was needed for quality film deposition,⁹ As it is seen in this study, when SiO_{2(s)} is removed primarily through mechanism A (Table V), a higher temperature process (e.g., a hydrogen bake) would be needed, whereas when it is mainly removed through mechanism B, no hydrogen bake might be required.

For 100% Ge growth, since GeO_{2(s)} is very unstable, the removal of GeO_{2(s)} is easily achieved at very low temperature (<50 °C) for all pressure and oxygen level studied.

Results from such thermodynamic analyses were utilized to improve the quality of epitaxial Si_{1-x}Ge_x films deposited in a conventional hot-wall tubular LPCVD reactor.¹⁷ The process involved the growth of a Si SEG buffer layer fol-

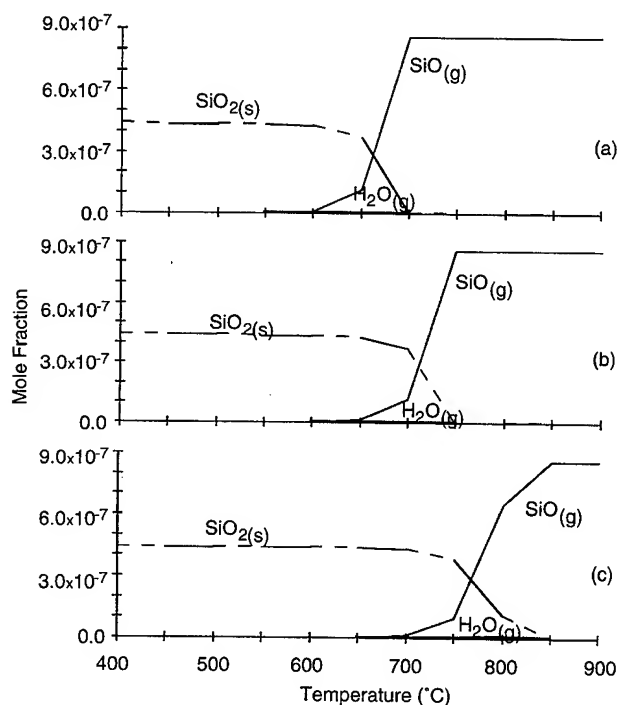


FIG. 1. Major oxygen-containing species for 0%–99% Ge in initial Si+Ge, water vapor level of 1.0 ppm, and (a) 0.1, (b) 1.0, and (c) 10.0 Torr. GeO_{2(g)} in these conditions exists in very small amounts (not shown).

TABLE V. Mechanisms of silicon dioxide removal from a substrate surface.

Mechanism A	Mechanism B
Removal of oxide by $\text{SiO}_{2(s)} + \text{Si}_{(s)} \text{ or } (g) \rightarrow 2\text{SiO}$	Removal of oxide by $\text{SiO}_{2(s)} + 2\text{H}_2 \rightarrow 2\text{H}_2\text{O} + \text{Si}_{(s)}$
SiO formation enhanced by decreasing pressure	H ₂ O formation enhanced by increasing pressure
An overlap of the SiO _{2(s)} and SiO profiles (Fig. 1)	A gap between the SiO _{2(s)} and SiO profiles (Fig. 2)
Favored at higher oxygen level	Favored at very low oxygen level and low level of Si gaseous sources

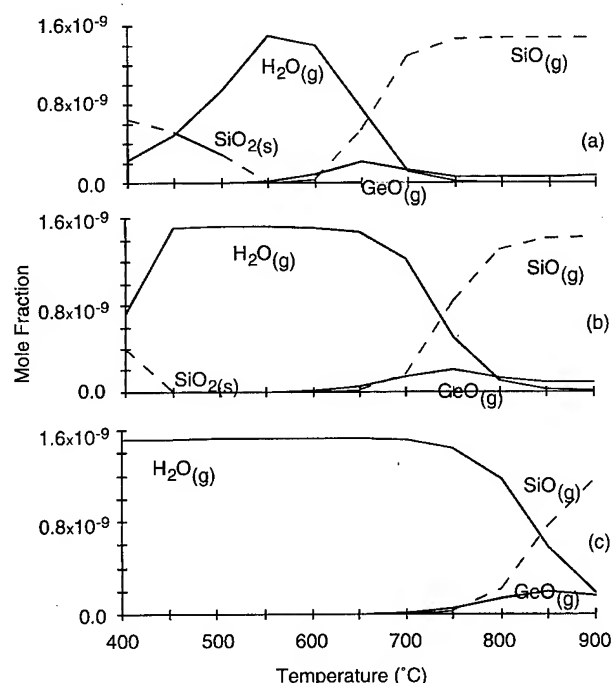


FIG. 2. Major oxygen-containing species for 99.5% Ge in initial Si+Ge, 0.002 ppm water vapor, and (a) 0.1, (b) 1.0, and (c) 10.0 Torr.

lowed by a $\text{Si}_{1-x}\text{Ge}_x$ SEG film. Whereas the deposition of the Si SEG buffer layer at 900 °C resulted in a good quality film indeed,¹⁷ oxide formation on the growth surface was found to be a problem during the deposition of $\text{Si}_{1-x}\text{Ge}_x$ films below 750 °C. A plausible explanation appeared to be that decreasing the deposition temperature from 900 °C to below 750 °C, the reactor environment changed from oxide-free to oxide-forming (e.g., Table III). Therefore, in order to keep the reactor environment oxide-free, a small flow of SiH_2Cl_2 was introduced during the temperature ramp down from 900 ° to the desired SiGe SEG temperature (usually less than 750 °C). The addition of SiH_2Cl_2 provided a Si source for the oxide removal through mechanism A (Table V), thus yielding an oxide-free surface for the epitaxial growth of $\text{Si}_{1-x}\text{Ge}_x$ films.¹⁵

B. Chlorine effect

The etch/deposition boundaries for pure Si and pure Ge show two different trends as illustrated by the contour plot in Fig. 3, where the lowest chlorine level to achieve etch (or the highest chlorine level at which deposition may take place) is shown through the parameter Cl/H with values between 0.075 and 0.11 at those conditions. For pure Si, the critical chlorine level increases with increasing system temperature, while pure Ge shows exactly the opposite trend. A possible explanation is that the concentrations of silicon-containing species (e.g., SiCl_4 , SiHCl_3) decrease with increasing temperature, while those of germanium (e.g., GeCl , GeCl_2) increase with increasing temperature.¹⁰ Therefore, at low temperatures, it takes more chlorine to achieve an etching favorable condition for Si than for Ge and vice versa at high

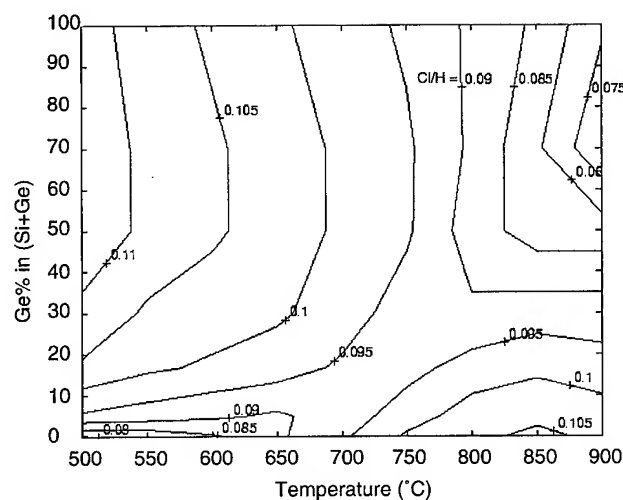


FIG. 3. Contour plot of the lowest chlorine level (Cl/H) to achieve etch (or the highest chlorine level at which deposition may take place) for $P = 1.0$ Torr and $(\text{Si}+\text{Ge})/\text{H} = 1:50$.

temperatures. Calculated results in this study for pure Si are in good agreement with experimental results in the literature, where the higher the temperature, the more chlorine is needed to achieve etch-favorable conditions.^{18–20} When a mixture of Si and Ge is used, the result becomes more intriguing. A saddle-type contour is obtained, because the conditions for two of the four corners require the highest critical chlorine level (the upper left and lower right corners of Fig. 3), while the other two require the lowest.

The dominant chlorine-containing species are SiCl_4 , GeCl_2 , and HCl as observed from the results. Their relative amounts at different temperatures provide a clue to the explanation of the observed trends. At a low temperature, the formation of SiCl_4 is the most favored, while GeCl_2 is the least favored. Therefore, as shown in Fig. 4, it takes more chlorine to achieve etching conditions for 16.7% Ge (in Si+Ge) than for pure Si, because chlorine preferentially tends to form SiCl_4 and HCl . At a high temperature, the process is reversed; GeCl_2 formation is now favored followed by HCl , therefore, less chlorine is now needed to achieve etch-favoring conditions for the 16.7% Ge case than for pure Si.

The effect of pressure on the etch/deposition boundary based on thermodynamic analyses is illustrated in Figs. 3, 5, and 6. The calculations have been performed at process pressures ranging from 0.1 to 10.0 Torr. Basically, as the system pressure increases, less chlorine (shown in the form of the parameter Cl/H in these figures) is needed to reach the etch/deposition boundary. This result is in agreement with the theoretical study by Chung *et al.* for the Si–H–Cl–O systems¹² as well as with the trend observed in the experimental study of Borland and Drowley at 25–80 Torr.¹⁷

Thermodynamic analyses like these can indeed provide effective and useful information for CVD process. Results such as the ones presented here could be essential for a basic understanding of these processes, which in turn can be useful

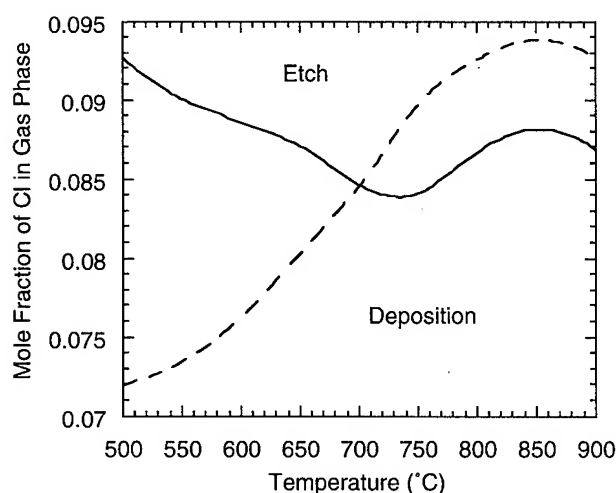


FIG. 4. The etch/deposition boundary for pure Si (dashed line) and 16.7% initial Ge (solid line) at $P=0.1$ Torr. The area above a line is the etch-favoring parameter region; the area below a line is the deposition-favoring region.

in experimental studies as well as related reactor design. Hence, in a very short period of time, such analyses coupled with a minimum number of experiments can result in optimized processing conditions and/or novel processing based on fundamental-knowledge-driven process property relationships (e.g., see Refs. 14 and 17).

IV. SUMMARY AND CONCLUSION

The effects of water vapor and chlorine on the epitaxial growth of $\text{Si}_{1-x}\text{Ge}_x$ films have been investigated. The required levels of water vapor (and oxygen) for oxide-free $\text{Si}_{1-x}\text{Ge}_x$ epitaxial growth are found to be similar to those for Si epitaxial growth for most ratios of the silicon to germanium gaseous sources. This appears to be a result of the

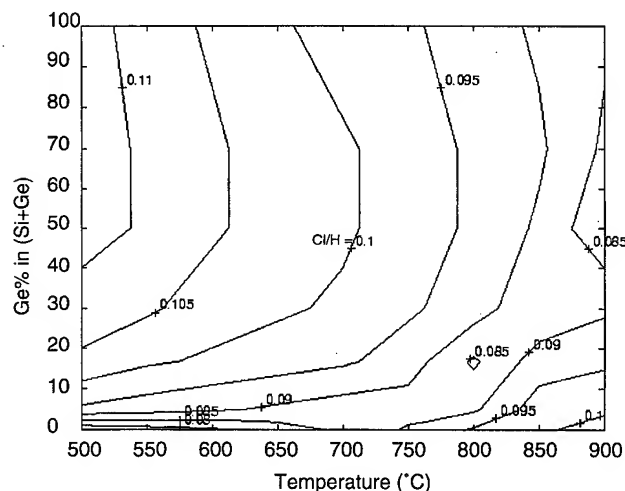


FIG. 5. Contour plot of the lowest chlorine level (Cl/H) to achieve etch (or the highest chlorine level that deposition may take place) for $P=0.1$ Torr and $(\text{Si}+\text{Ge})\text{:H}=1\text{:}50$.

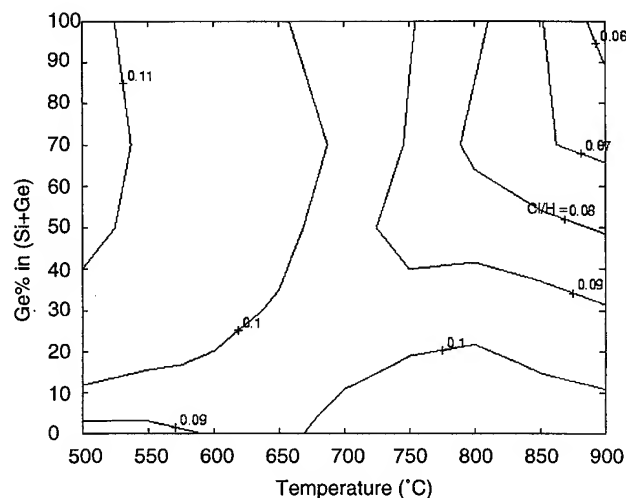


FIG. 6. Contour plot of the lowest chlorine level (Cl/H) to achieve etch (or the highest chlorine level at which deposition may take place) for $P=10.0$ Torr and $(\text{Si}+\text{Ge})\text{:H}=1\text{:}50$.

fact that germanium does not form major oxygen-containing species. At high excesses of the Ge gaseous source and low water vapor levels, however, $\text{SiO}_{2(s)}$ is removed through reaction with H_2 to form gas phase H_2O instead of reacting with Si to form SiO . Hence, at large excesses of Ge gas phase sources and low system oxygen level, the critical temperature above which oxide-free epitaxial growth occurs decreases with increasing system pressure. Calculated results are in good agreement with available experimental results for Si epitaxial growth.

Chlorine concentrations needed for $\text{Si}_{1-x}\text{Ge}_x$ etching are found to decrease with increasing pressure. The etch/deposition boundaries for pure Si and pure Ge show opposite trends with respect to increasing chlorine concentration in the system, and this can be explained by their easiness to form chlorinated species. It is found that the formation of SiCl_4 is favored over GeCl_2 at low temperatures, while the reverse is true at high temperatures. Therefore, the etch/deposition boundary for $\text{Si}_{1-x}\text{Ge}_x$ forms a saddlelike contour. Again, calculated results are in good agreement with available experimental results from Si epitaxial growth.

ACKNOWLEDGMENTS

The authors are grateful to Dr. B. J. McBride for parts of the thermodynamic data used. Financial support by the National Science Foundation is also gratefully acknowledged.

- ¹R. Hull, J. C. Bean, D. J. Werder, and R. E. Leibenguth, *Appl. Phys. Lett.* **52**, 1605 (1988).
- ²R. H. M. van de Leur, A. J. G. Schellingerhout, F. Tuinstra, and J. E. Mooij, *J. Appl. Phys.* **64**, 3043 (1988).
- ³M. Racanelli and D. W. Greve, *J. Occup. Med.* **32** (1991).
- ⁴F. W. Smith and G. Ghidini, *J. Electrochem. Soc.* **129**, 1300 (1982).
- ⁵G. Ghidini and F. W. Smith, *J. Electrochem. Soc.* **131**, 2924 (1984).
- ⁶B. S. Meyerson, *J. Appl. Phys.* **48**, 797 (1986).
- ⁷J. A. Friedrich, G. W. Neudeck, and S. T. Liu, *Appl. Phys. Lett.* **53**, 2543 (1988).

- ⁸M. R. Goulding, *Mater. Sci. Eng. B* **17**, 47 (1993).
- ⁹B. S. Meyerson, K. J. Uram, and F. K. LeGoues, *Appl. Phys. Lett.* **53**, 2555 (1988).
- ¹⁰I.-M. Lee and C. G. Takoudis, *J. Electrochem. Soc.* **143**, 719 (1996).
- ¹¹N. M. Hwang and D. Y. Yoon, *J. Cryst. Growth* **143**, 103 (1994).
- ¹²C.-H. Chung, S.-W. Rhee, and S. H. Moon, *J. Electrochem. Soc.* **142**, 2405 (1995).
- ¹³S. Gordon and B. J. McBride, Technical Report SP-273, NASA, NASA Lewis Research Center (1971).
- ¹⁴W.-C. Wang, J. P. Denton, G. W. Neudeck, I.-M. Lee, C. G. Takoudis, M. T. K. Koh, and E. P. Kvam, *J. Vac. Sci. Technol. B* **15**, 138 (1997).
- ¹⁵(a) W. H. Gaynor, research thesis, School of Chemical Engineering, Purdue University, 1989; (b) W. H. Gaynor, G. W. Neudeck, and C. G. Takoudis, *J. Vac. Sci. Technol. A* **14**, 3224 (1996).
- ¹⁶D. W. Greve and M. Racanelli, *J. Vac. Sci. Technol. B* **8**, 511 (1990).
- ¹⁷I.-M. Lee, C. G. Takoudis, W.-C. Wang, J. P. Denton, G. W. Neudeck, M. T. K. Koh, and E. P. Kvam, *J. Electrochem. Soc.* **144**, 1095 (1997).
- ¹⁸J. O. Borland and C. I. Drowley, *Solid State Technol.* Aug. 141 (1985).
- ¹⁹M. C. Arst, K. N. Ritz, S. Redkar, J. O. Borland, and J. Hann, *J. Mater. Res.* **6**, 784 (1991).
- ²⁰K. E. Violette, P. A. O'Neil, M. C. Ozturk, K. Christensen, and D. M. Maher, *J. Electrochem. Soc.* **96-5**, 374 (1996).

Integration issues for 850 nm optical modulators on Si electronics by direct epitaxy

J. E. Cunningham and W. Y. Jan

Bell Laboratories, Lucent Technologies, Holmdel, New Jersey

(Received 14 April 1997; accepted 18 April 1997)

We report on the various strategies for integrating III-V semiconductors to very large scale integrated Si logic along with their strengths and limitations. We also detail the monolithic integration method involving direct epitaxial growth where several material science issues have arisen. They require fundamental investigation to further advance implementation. Those issues are (i) the need for a highly ordered array of bilayer steps as an initial Si surface condition for heteroepitaxy, (ii) lower Si oxide desorption temperature, e.g., via electron cyclotron resonance plasma treatments to preserve the bilayer step ordering, and (iii) thinner strain relief layers to reduce the III-V on Si topology mismatch. © 1997 American Vacuum Society.
[S0734-211X(97)06404-4]

I. INTRODUCTION

It has long been recognized that once III-V devices are integrated to Si integrated circuits (ICs) then new and unique systems and/or applications would be enabled. Specific examples of applications targeted to benefit from this integration are optical interconnects,¹ switching,² infrared focal plane arrays,³ and high performance Avalanche Photodiodes.⁴ To date, significant advances in nearly all the above categories have been demonstrated as evidenced by the emergence of optoelectronic chips as either products or elements in system test beds. One of the most striking achievements in integration technologies has been through a process of hybrid bump bonding III-V to Si ICs with metallic solder.⁵ Two other important integration technologies are direct monolithic growth and contact bonding by interfacial fusion (of the two crystals).

The primary factors driving the marriage between Si and III-V materials have been applications where large amounts of information flow on and off an electronic chip are required. For situations that implement high bandwidth and dense interconnections, optical interconnects to the electronic chip appear to be a sensible solution. For sufficiently high bandwidths they may become the only solution. An example of this is shown in Fig. 1 where the computational capability is enhanced by optics over a pure electronic interconnect solution as projected by Krishnamoorthy and Miller.⁶ The data was taken from the Semiconductor Industry Association (SIA) roadmap.⁷ When information flow onto the chip is taken into account, then the processing capacity of Si electronics is given by the left ordinate. Diamonds give the case when the input is through interconnects on Si. Interconnect characteristics here are deduced from SIA predictions on future capabilities. On the other hand, optical interconnects, as indicated by solid circles, do enhance the capacity of Si electronics since the interconnect limitation has been overcome. The data was deduced from Ref. 6. Figure 1 shows optical interconnects become increasingly important as the feature size shrinks.

But optoelectronic systems also add complexity to appli-

cations both from a technology and a system viewpoint. On technological grounds optoelectronic chips tend to be small 64×64 element arrays, e.g., for flip chip bonded quantum well modulators to Si. Here, chips of size 1 cm have been demonstrated. For focal plane arrays, 640×480 , arrays (TV pixelization) have been accomplished by hybrid bonding resulting in 1 in. size chips. Hence, wafer scale integration is not yet available for the hybrid optoelectronic solution and low cost fabrication is not yet an option. On the other hand wafer scale integration remains possible for both fusion bonding and the monolithic integration methods. In fusion bonding Si and III-V wafers are joined physically by high temperature/pressure treatments. However, fusion bonding appears appropriate when the physical heterojunction between Si and a III-V provide a distinct advantage such as in avalanche photo-diode, applications. It is less attractive for optoelectronic chips based on very large scale integrated Si. Monolithic integration is amenable to wafer scale integration. Nonetheless, it is limited by several fundamental material science problems which the subject matter of the remainder of this article.

II. MONOLITHIC GROWTH ISSUES

Recently, we reported the first successful monolithic integration of GaAs multiple quantum well (MQW) light modulators grown directly on Si and subsequently interconnected electrically to pre-existing Si ICs.⁸ The method of process we employ to perform that integration is shown in Fig. 2. From a growth viewpoint several issues arise. First, to facilitate integration miscut Si (100) wafers were used under the presumption that antiphase domain GaAs growth could be avoided.⁹ Recently, it has been reported that single domain GaAs growth could be obtained without the miscut.¹⁰ Further, the miscut could be detrimental to transistors in applications implementing more aggressive IC densities due to their thinner oxide gate thickness. Also, the misorientation can produce internal piezoelectric fields in the MQW from incomplete strain relief and thereby complicate the quantum confined Stark effect (QCSE) that is needed for the modula-

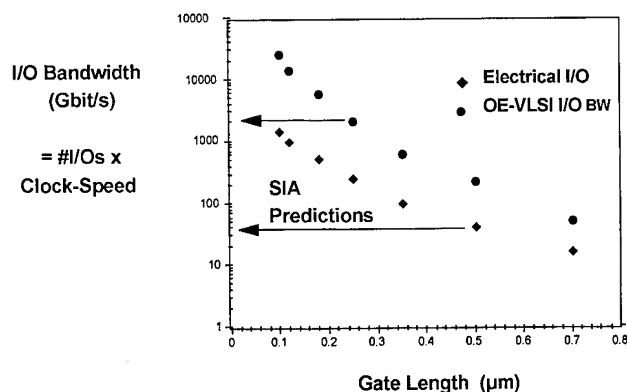


FIG. 1. Projections for the computational power of Si with and without optical interconnects for different gate length. Diamonds give the processing capability of Si electronics when utilizing optical input and output. The solid circles give the result when limited by the interconnect characteristics that the SIA roadmap projects.

tor functionality.¹¹ Second, the high temperature Si oxide desorption step is comparable to the total thermal budget of typical Si IC processing steps. Thus, the oxide desorption step in the integration process could degrade Si ICs. One further consideration is that the ICs cannot have the final Al metallization when implementing the processes within Fig. 2 since they cannot withstand temperatures above 450 °C. Although Si foundries do provide ICs without final Al metal, the task to remetalize after III–V growth is very difficult, particularly when the gate feature size shrinks.

To address these issues we compare MQW on different miscut Si wafers and show that the miscut is not only required, but also the staircase must exhibit a very high degree of bilayer step ordering to implement successful integration. In efforts to reduce the oxide desorption temperature, we use electron cyclotron resonance (ECR) methods to proton implant the shiriki oxide on Si. This process differs from the direct ECR assisted oxide desorption process previously reported.¹² In addition, we investigate Si step ordering characteristics under the different ECR treatments. Finally, we demonstrate the monolithic approach can support high quality shallow quantum well modulators ($x=0.025$).

III. GaAs EPITAXY ON Si

GaAs on Si epitaxy was performed using a modified two-step process reported by Nishi.¹³ Prior to deposition careful examination of the quality of bilayer steps was analyzed by reflection high-energy electron diffraction (RHEED) investigations in a Si desorption chamber.^{14,15} For the nucleation stage both low growth temperature (400–450 °C) and low growth rate (0.2–0.5 $\mu\text{m/h}$) were used. The V/III flux was maintained close to 1.0 during all stages of epitaxy except at $T_g = 640$ °C (the end of the second step process for the sacrificial buffer growth) where the ratio increased to 1.5. The overall quality of the ensuing MQW epitaxy could be judged in the nucleation stage of growth. Here, the first monolayer of GaAs deposition must show a streaked pattern in RHEED. Upon 10–100 Å GaAs deposition on Si the RHEED trans-

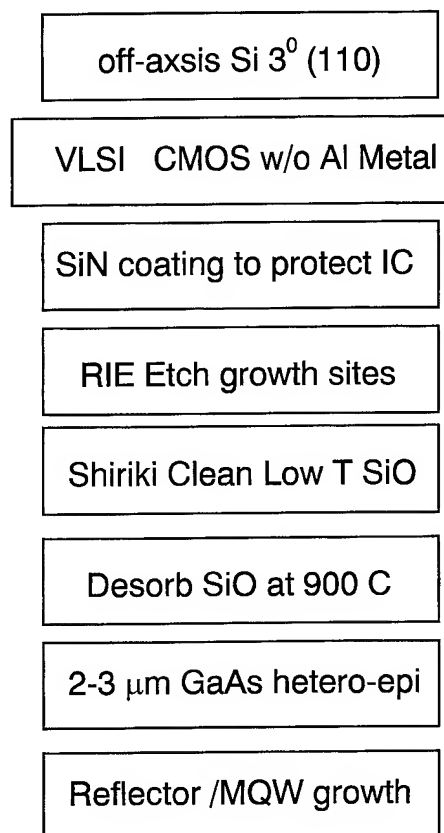


FIG. 2. Process for implementing monolithic integration to Si by epitaxial growth.

forms to elongated spots which become increasingly spotty with thickness for the case of the incident beam aligned perpendicular to the staircase (along the step ledge). On the other hand, for the beam directed along the staircase, RHEED continued to show broken streaks with both the 0th and 1st order Laue Zones visible. This suggests that GaAs growth proceeds by filling in the staircase as flat rectangular blocks aligned lengthwise to the step edges. Such a growth mode differs from the Stranski–Krastanov (SK) growth mode¹⁶ widely believed to apply to GaAs on Si. In SK growth three-dimensional nucleation occurs randomly on the surface. A scanning electron microscope (SEM) photograph illustrating the growth mode found on Si is shown in Fig. 3(a) and was obtained on 50 Å of GaAs deposition on Si below a thin As cap. The photograph shows highly elongated GaAs domains that have preferentially ordered along the staircase. When the 1st order Laue Zone in RHEED cannot be detected, then the top surface of the GaAs domains are not flat and a more SK-like growth mode applies. An example of this is shown in the SEM photo in Fig. 3(b) where random nucleation with respect to the staircase occurs. Again a thin 50 Å GaAs layer was deposited. Notice also that growth proceeds close to a hemispherical ball shape much like SK predictions. When epitaxy proceeds in this way then, very poor quality modulators ensue. Several factors disorder the bilayer staircase of steps that lead to poor epitaxy of the type shown in Fig. 3(b). These have been identified in Ref. 14.

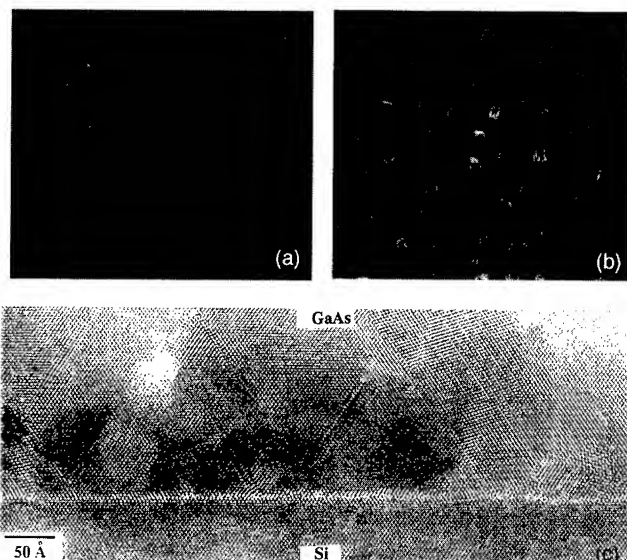


FIG. 3. (a) SEM from 50 Å GaAs on Si. The staircase is along the bar. The bar length is 1500 Å. (b) SEM from 50 Å GaAs on Si in which the staircase contained disorder. (c) Atomic resolution TEM from GaAs on Si in which initial epitaxy followed (a).

Upon reaching the second stage of GaAs growth, (2×4) GaAs surface reconstructions were observed for both on axis and miscut surfaces, provided the initial stage epitaxy proceeded correctly.

An important problem associated with GaAs on Si heteroepitaxy is the $3 \mu\text{m}$ strain relief layer that is needed before actual device quality material can be grown. This leads to large topology features that make further processing of integrated structures difficult (such as the gate metal vias interconnections). It would be desirable to reduce the epilayer deposition and in fact to employ an integral Bragg reflector ($1.5 \mu\text{m} = 95\%$ reflectivity) as the strain relief layer, provided the surface morphology is sufficiently smooth to avoid scattering losses in the mirror. The fundamental requirement of the thick strain relief layer is not clear, but much of our investigations reveal that stacking faults from antiphase domains are a principal problem. For example, even when GaAs epitaxy on Si follows that in Fig. 3(a) the primary defect in the second stage growth is stacking faults. This is clearly evident from the atomic resolution transmission electron microscopy (TEM) view shown in Fig. 2(c). Insight into highly efficient strain relief processes can be gained from the heteroepitaxial case of InP on GaAs where a 4% lattice mismatch without antiphase boundaries occurs. By using a two-stage growth process for on axis substrates we find that excellent quantum well formation can be detected after only 1000 Å of InP deposition on GaAs.¹⁷ Also the surface morphology is exceptionally smooth and surface reconstructions distinct. During the initial stage of epitaxy, both GaAs on Si and InP on GaAs produce similar RHEED patterns after 100 Å deposition. It is the second stage epitaxy where differences develop with InP on GaAs recovery being extremely rapid with deposition but GaAs on Si comparatively slow.

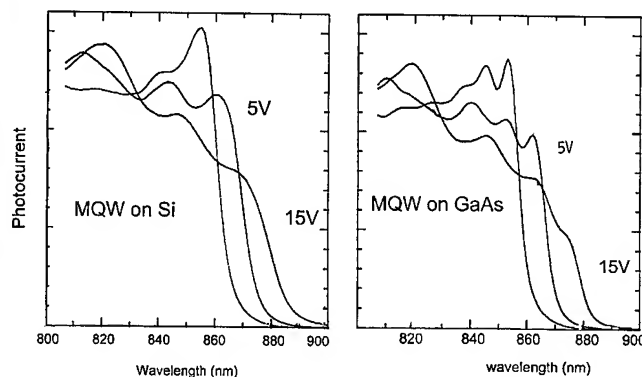


FIG. 4. Photocurrent from MQW on miscut Si and GaAs substrates.

IV. OPTICAL MEASUREMENTS

MQW modulator performance was investigated with regard to excitonic quality and (QCSE) characteristics by using photocurrent spectroscopy on structures grown under different Si surface treatments prior to the heteroepitaxy growth. Figure 4 compares photocurrent spectroscopy for 850 nm MQW modulators on miscut Si (heteroepitaxy) with the identical structure on GaAs (homoepitaxy). Here, the growth of an identical modulator was accomplished by mounting a $3 \mu\text{m}$ thick GaAs on a Si epilayer adjacent to a GaAs substrate prior to the MQW modulator deposition. The MQW structure consisted of 50 periods of AlGaAs/GaAs with barriers of ($x=3$) 50 Å thick and a 95 Å wide quantum well. Judging the strength of photocurrent signal on each. 0 V spectra, we conclude that $2d$ excitonic formation¹⁸ on Si is comparable to that on GaAs. Also, QCSE is virtually identical for the two systems which further implies that quantum well interfaces have low disorder and few defects. The small differences between the two are in the position of the excitonic resonance (850 versus 843 nm), the light and heavy hole splitting (8 versus 0 meV) and the half-width at half-maximum (10 versus 6 meV), respectively, for (homo versus heteroepitaxy). The differences can be explained by assuming residual tensile strain exists in the MQW layers. This strain is expected because of the thermal expansion mismatch between GaAs and Si for MQW growth at 600 °C.

Figure 5(a) shows spectra taken from MQW modulators grown either on axis or on a surface misoriented by 2° and measured at room temperature. Each structure was grown at 600 °C atop a $2\text{--}3 \mu\text{m}$ thick GaAs buffer layer on Si. Figure 5(a) shows photocurrent dependence from a shallow MQW ($x=0.025$) that exhibits QCSE under applied field. In fact, QCSE quality on Si is comparable to those grown on GaAs. At 0 V, the excitonic linewidth is 9 meV full width at half-maximum (FWHM) which is identical to the best obtained in the GaAs/AlGaAs system.¹⁹ These results are quite surprising since the low barriers can only weakly confine excitons and QCSE is easily destroyed by sample imperfection. The results demonstrate that MQW modulators on Si can be robust.

Figure 5(b) shows little to no QCSE for MQW grown directly on axis Si. Instead, the edge broadens much like

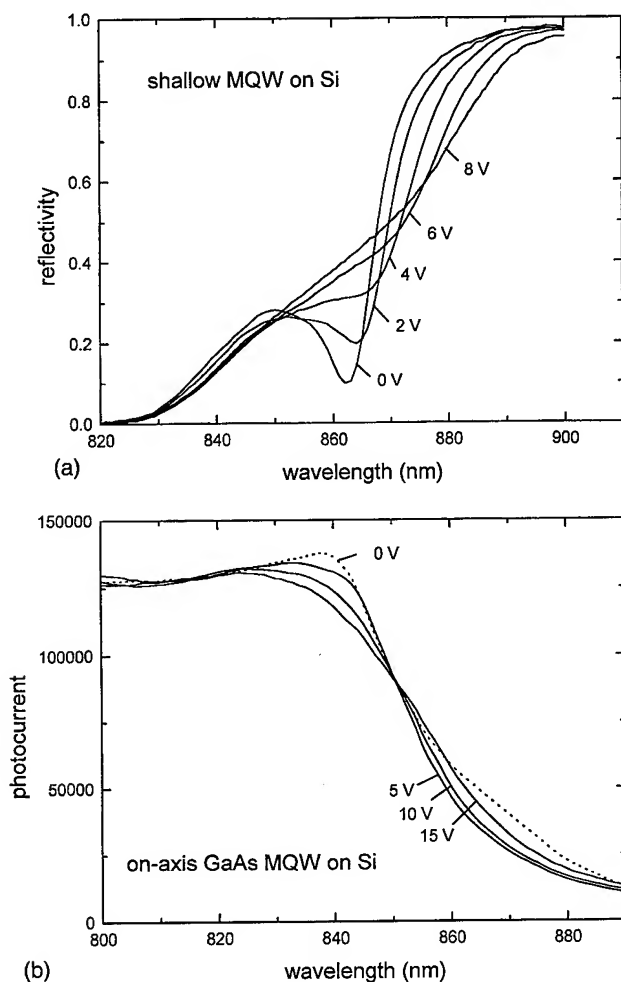


FIG. 5. Photocurrent vs wavelength of MQW on (a) on-axis growth and (b) 2° miscut.

bulk GaAs. The surface morphology for this case was excellent and the x-ray FWHM was 500 arcsecs. No antiphase domains were detected in RHEED. Nevertheless, the position of the absorption edge is found to be blue shifted from bulk GaAs because significant size quantization occurs. This is supported by transmission electron microscopy (TEM) which resolved chemical contrast from the barrier ($x = 0.3$)/well of the MQW. Note the FWHM of the exciton for Fig. 5(b) is 20 meV and therefore much wider than bulk GaAs, 11 meV. It is therefore plausible that substantial internal fields from various sources of defects occur for on axis growth. Most defects appear to be stacking fault related. Nevertheless, the necessity of bilayer steps for the integration is clearly established. We have further correlated that detectable surface roughness (deduced under phase contrast optical microscopy) destroys the QCSE in modulators and that high crystallinity is less important.

V. ECR FOR REDUCED OXIDE DESORPTION TEMPERATURES

Our investigations of oxide desorption under assistance by atomic H reveal that considerable damage can occur to a

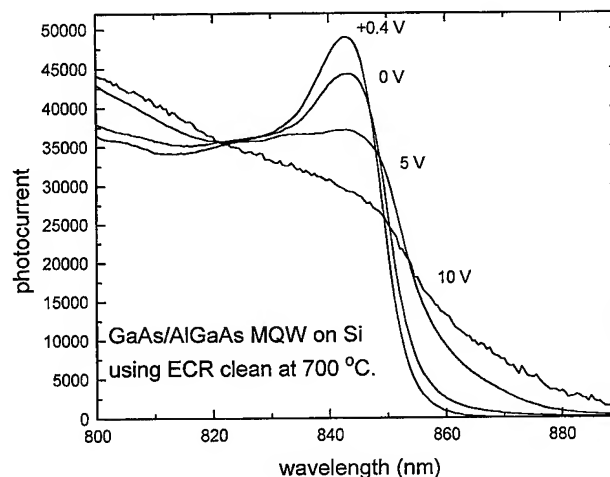


FIG. 6. Photocurrent vs wavelength of MQW on ECR-exposed, miscut Si.

highly ordered Si step staircase. All our applications of direct ECR assistance in the desorption process damaged the staircase with most conditions producing faceting. This can be understood as follows. First, a bilayer staircase is thermodynamically unstable with temperature upon H adsorption at the ledges and secondly atomic H plasma etches miscut Si anisotropically when silane forms. In addition, monohydride and dihydride formation passivates Si dangling bonds that would otherwise minimize surface free energy for the bilayer step configuration through the benefit of dimerization energy. Much like the other disordered surfaces previously described, the ECR desorbed surfaces lead to poor GaAs heteroepitaxy. As an alternative, we investigate a different approach based upon atomic H implantation into the Si/SiO₂ system prior to the oxide desorption step. Auger spectra were before and after SiO₂ exposure to atomic H for 15 min at a temperature of 300 °C. They reveal the oxide layer is still present even after atomic H exposure. In fact, significant oxide removal does not commence until 60 min of atomic H exposure. Even after 90 min of ECR exposure faint traces of oxide remain. Upon heating the sample after a 10 min ECR exposure the oxide is then found to desorb at 730 °C, i.e., 100 °C lower than unexposed samples. Bilayer step ordering of Si, after undergoing the latter process, showed spot splittings in RHEED (from order steps) nearly comparable in quality to those undergoing high temperature desorption. GaAs heteroepitaxy was successfully performed on these surfaces. MQW fabricated on these surfaces.

Figure 6 is the modulation performance of MQW grown on off axis Si containing the ECR implantation prior to low temperature oxide desorption. The exciton has FWHM of 11 meV and is thus only slightly broader than MQW on GaAs. QCSE does occur in this sample but is noticeably smaller (smaller shifts with applied fields). In addition, for fields above 10^5 V cm⁻¹ QCSE changes character to exhibit bulk-like behavior. The origin of the QCSE transition from 2d to 3d behavior is most likely defect induced. (Excitonic absorption that red shifts under bias is characteristic of QW struc-

tures which are $2d$ in origin, bulk layer do not red shift and hence are characteristic of $3d$ excitons).

VI. CONCLUSION

We have shown that growth of GaAs modulators on Si for interconnect applications requires a very high quality Si surface in terms of step ordering characteristics and cleanliness for heteroepitaxy. ECR implantation into SiO_2 can reduce oxide desorption temperatures and still permit observation of both bilayer surface step ordering and QCSE from MQW. However, there remains much research to solve the monolithic integration approach. Specifically, methods must be developed to (i) desorb the oxide below 450°C while retaining a high degree of step ordering, (ii) develop a staircase of bilayer steps by selective etching into the on axis Si(100), (iii) reduce the topology differences with improved strain relief processes.

Presented at the 43rd National Symposium of the American Vacuum Society, Philadelphia, PA, 12–18 October 1996.

¹M. J. Goodwin, A. J. Moseley, M. Q. Kearley, R. C. Morris, C. J. G. Kirkby, J. Thomson, R. C. Googfellow, and I. Bennion, *J. Lightwave Technol.* **9**, 1639 (1991).

²A. L. Lentine, K. W. Goossen, J. A. Walker, L. M. F. Chirovsky, L. A. D'Asaro, S. P. Hui, B. Tseng, R. Leibenguth, J. E. Cunningham, W. Y. Jan, J. M. Kuo, D. Dahringer, D. Kossives, D. Bacon, G. Livescue, R. L. Morrison, R. A. Novotny, and D. B. Buchholz, *Conference on Laser and Electro-optics (CLEO-96)* May 1996, Vol. 9, pp. 517, 518.

³*Infrared Imaging News* 2, 3 (1996).

⁴A. R. Hawkins, T. E. Reynolds, D. R. England, D. I. Babic, M. J. Mon-

dry, K. Strebuel, and J. E. Bowers, *Appl. Phys. Lett.* **68**, 3692 (1996).

⁵K. W. Goossen, J. E. Cunningham, and W. Y. Jan, *IEEE Photonics Technol. Lett.* **5**, 776 (1993).

⁶A. V. Krishnamoorthy and D. A. B. Miller, *Proceedings of the Lasers Electrooptic Society Annual Meeting, Optical Interconnects and Processing Systems*, San Francisco, CA, October 1996.

⁷National Technology Roadmap for Semiconductors (Semiconductor Industry Association, 1994), p. B2.

⁸K. W. Goossen, J. A. Walker, J. E. Cunningham, W. Y. Jan, and D. A. B. Miller, *OSA Proceedings on Photonics in Switching*, edited by J. W. Goodman and R. C. Alfarness, CA, 1993 (unpublished), Vol. 16, p. 50.

⁹R. P. Gale, B. Y. Tsaur, J. C. C. Fan, F. M. Davis, and G. W. Turner, *Proceedings of the Sixteenth Photovoltaic Specialist Conference*, 1981 (unpublished), p. 1051.

¹⁰W. I. Wang, *Appl. Phys. Lett.* **44**, 1149 (1984).

¹¹J. E. Cunningham, K. W. Goossen, and W. Jan, *J. Cryst. Growth* **127**, 184 (1993).

¹²Y. Kunitsugu, I. Suemune, Y. Tanaka, Y. Kan, and M. Yamanishi, *J. Cryst. Growth* **95**, 91 (1989).

¹³S. Nishi, H. Inomata, M. Akiyama, and K. Kaminish, *Jpn. J. Appl. Phys.* **24**, 391 (1985).

¹⁴J. E. Cunningham, K. W. Goossen, J. A. Walker, W. Jan, M. Santosand, and D. A. B. Miller, *J. Vac. Sci. Technol. B* **12**, 1246 (1994).

¹⁵P. R. Purkate, J. M. van Hove, and P. I. Cohen, *Appl. Phys. Lett.* **44**, 456 (1984).

¹⁶J. N. Stranski and L. Krastanov, *Ber. Akad. Wiss. Wien*, **146**, 797 (1938).

¹⁷R. N. Pathak, J. E. Cunningham, W. Jan, W. K. Knox, S. Tsuda, and B. Collings (unpublished).

¹⁸Excitons exhibit widely different absorption characteristics in $2d$ vs $3d$ conditions. Oscillator strength at threshold is enhanced by a factor of 4 in $2d$. Under bias $2d$ excitons exhibit QCSE (red shifting) whereas $3d$ excitons broaden (Franz-Keldysh effect).

¹⁹K. W. Goossen, J. E. Cunningham, W. Jan, and J. A. Walker, *IEEE Photonics Technol. Lett.* **7**, 763 (1992).

Interfacial reaction behavior of Pt, Pd, and Ni on ZnSe

K. J. Duxstad and E. E. Haller^{a)}

Department of Materials Science and Mineral Engineering, University of California at Berkeley and Lawrence Berkeley National Laboratory, Berkeley, California 94720

K. M. Yu, E. D. Bourret, X. W. Lin,^{b)} S. Ruvimov, Z. Liliental-Weber, and J. Washburn
Lawrence Berkeley National Laboratory, Berkeley, California 94720

(Received 3 January 1997; accepted 4 April 1997)

The solid-state reactions of Pt, Pd, and Ni thin film contacts on ZnSe were studied. Pd has been observed to react at 200 °C to form a ternary, epitaxial phase, $\text{Pd}_{5+x}\text{ZnSe}$. This phase is stable up to 450 °C, when Pd begins to diffuse through the ZnSe into the GaAs substrate. Pt begins to react at 575 °C and forms a layer of Pt_5Se_4 at the Pt/ZnSe interface. Above the interfacial layer there is a Pt–Zn solid solution, but no Pt–Zn phases were observed. After annealing at 675 °C, the Pt_5Se_4 phase is no longer observed and Pt–Zn phases form. A small amount of interdiffusion at the Ni/ZnSe interface has been observed by transmission electron microscopy after annealing at 300 °C. Annealing at 425 °C results in the formation of laterally separated grains of a metastable Ni_xSe phase. After annealing at 450 °C, grains of NiSe are observed. The similarities between these reactions and the near-noble metal reactions on Si and GaAs are discussed as well as the possibility of using these reactions for forming electrical contacts to ZnSe. © 1997 American Vacuum Society. [S0734-211X(97)00704-X]

I. INTRODUCTION

ZnSe is a semiconductor with a wide direct band gap of 2.7 eV at room temperature. It has been used for the development of blue light emitting devices. Continuous-wave¹ and pulsed^{2,3} room-temperature operation of ZnSe lasers have been reported. For all device applications, one important consideration is the ability to form stable, low resistance contacts. For example, the performance of laser diodes could be improved by reducing the specific contact resistance to the *p*-type layers of the device.¹ Currently, the *p*-type contacts are Zn(Se,Te) graded layers. Ohmic contacts to *p*-type ZnSe have been very elusive because of the very deep valence band maximum and the inability to dope highly *p*-type. The valence band maximum is 6.4 eV below the vacuum level, and the largest work function metal, Pt, has a work function of only 5.3 eV. According to the Schottky model, the metals that will give the smallest barriers on *p*-type material are those with the largest work functions, Pt, Pd, and Ni. Pt has been shown to have a lower barrier height on *p*-type ZnSe than Au.⁴ Since no elemental metal is able to produce an ohmic contact on *p*-ZnSe, it is necessary to investigate more complicated contacting schemes.

The ability to engineer useful contacts requires an understanding of the thin film reactions that occur at the interface between the metal and semiconductor. Also, the characteristics of the contact must not change during processing or device operation. An understanding of the reaction behavior of metals on semiconductors has proven to be essential in the development of devices. Silicide contacts on silicon and the more recent development of the Ge/Pd based contacts on

GaAs formed by solid state regrowth are prominent examples.⁵ For a contact structure on ZnSe using the solid phase regrowth concept, the formation of a useful alloyed contact must occur at temperatures below 300 °C because the electrical characteristics of ZnSe are known to degrade with annealing above this temperature.^{6,7}

The near-noble metals are desirable for contacts because of their large work functions, their relatively high reactivity with semiconductors, and their resistance to oxidation and corrosion. In this article, we report on studies of the thermal stability of Pt, Pd, and Ni on ZnSe. The use of these metals as contacts requires the assurance that only controllable interfacial reactions will take place at the metal/semiconductor interface.

Pt, Pd, and Ni contacts have been studied extensively on Si and GaAs.^{8–10} Their behavior on these two semiconductors has been well characterized. On both Si and GaAs, Pd is known to react at room temperature whereas Ni reacts at 200 °C.^{8,9} On GaAs both Pd and Ni first form polycrystalline ternary phases which grow epitaxially on the substrate. Pt, however, reacts at 250 °C on GaAs to form vertically separated metal (M)–Ga and M–As phases.¹⁰ The final stable phases of all of the M–GaAs reactions are binary M_xGa_y and M_xAs_y phases.

In a previous paper,¹¹ we reported the formation of the $\text{Pd}_{5+x}\text{ZnSe}$ phase as the product of the initial reaction between Pd and ZnSe. In this article, we discuss the stability range of this phase as well as the reaction behavior of Pt and Ni. Pd is expected to react at the lowest temperature on ZnSe since it has the lowest binding energy of the three metals and reacts at the lowest temperature on Si and GaAs. A photoemission study of the stability of Pd on ZnSe showed that Zn dissolves into the metal overlayer. We have observed simi-

^{a)}Electronic mail: eehaller@lbl.gov

^{b)}Present address: VLSI Technology, Inc., 1109 McKay Drive, San Jose, CA 95131.

larities between the reactions of the near-noble metals on GaAs and on ZnSe.

II. EXPERIMENT

The ZnSe films used in this study were undoped and grown by metalorganic vapor phase epitaxy on semi-insulating GaAs.¹² The films were 0.5–0.9 μm thick. Before depositing the metal, the samples were solvent cleaned in a sequence of trichloroethane, acetone, and methanol. The samples were then etched for 30 s in HF, rinsed in distilled H_2O , and blown dry with N_2 gas. They were immediately loaded into the deposition chamber of an electron beam evaporator or a radio-frequency sputtering system. The metal films were evaporated or sputtered to a thickness between 300 and 800 Å. In this study, we report on the reactions of sputtered metal films on ZnSe. The reactions of evaporated films of Pd and Ni were the same as those for the sputtered films; however to obtain a consistent reaction of Pt on ZnSe the surface had to be backspattered prior to metal deposition. Backsputtering was done with Ar^+ ions at a rate of 1 Å/s. Samples were annealed in flowing N_2 for times ranging between 5 and 60 min.

Changes in the composition with depth were characterized using Rutherford backscattering spectrometry (RBS). RBS has a depth resolution between 50 and 100 Å, indicating that interdiffusion phenomena on a smaller scale cannot be analyzed. RBS was performed with a 1.95 MeV $^4\text{He}^+$ beam at a scattering angle of 165° . New phase formation was determined using x-ray diffraction (XRD). XRD was performed on a Siemens D5000 diffractometer in both Bragg-Brentano (θ - 2θ) and glancing incident angle geometries. Transmission electron microscopy (TEM) was used to observe the morphology of the interfacial reaction. Selected area diffraction (SAD) provided confirmation of new phase formation and determined any orientation relationship of the new phase with the substrate. Energy dispersive spectroscopy (EDS) was done in the TEM to determine the composition of the precipitates. TEM was performed employing a JEOL 200CX high-resolution electron microscope. Cross-sectional samples were prepared by gluing two 1-mm-thick pieces face to face, followed by mechanical grinding, and finally ion milling on a liquid nitrogen cooled stage to obtain electron transparency.

III. RESULTS

A. Pd/ZnSe

In a previous paper, we reported the initial reaction behavior of Pd on ZnSe.¹¹ After reiterating the major findings, we will discuss the phase stability of the Pd/ZnSe system up to temperatures of 500 °C. We have shown that Pd first begins to react during annealing at 200 °C for 10 min. The reaction results in the formation of $\text{Pd}_{5+x}\text{ZnSe}$, $0 < x < 1$. This phase was previously observed in a bulk reaction,¹³ and was determined to be a tetragonal phase with $a = 3.952$ Å and $c = 6.914$ Å. Identification of the phase formed in the thin film reaction was made by XRD and SAD. Figure 1

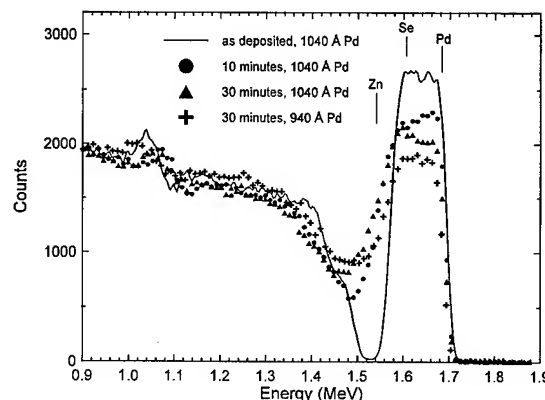


FIG. 1. Rutherford backscattering spectra of Pd on ZnSe annealed for various times at 200 °C with different as-deposited thicknesses.

shows RBS spectra from samples annealed at 200 °C for 10 and 30 min compared to an as-deposited sample. The spectral feature in Fig. 1 at approximately 1 MeV recoil energy marks the interface between the ZnSe and GaAs. The sample with an initial Pd thickness of 1040 Å has not reacted completely after annealing for 30 min at 200 °C. The ternary phase has formed, but with the stoichiometry of Pd_6ZnSe . Conversely, the film with an initial thickness of 940 Å has reacted completely after annealing for 30 min and the stoichiometry of the ternary phase is Pd_5ZnSe . Further annealing does not change the stoichiometry of this ternary phase significantly. The thickness of the ternary layer is approximately 120 nm. After annealing at 300 and 400 °C the depth of the reaction increases slightly and the XRD peaks shift to slightly smaller d values indicating a decrease in the measured value of c from 6.92 to 6.88 Å. This may be due to a slight change in the composition of the layer.

Figure 2 shows corresponding cross-sectional TEM micrographs of the samples shown in Fig. 1. In the reacted samples, the ternary phase forms a complete layer. The interface in the annealed sample is smoother than that in the as-deposited sample. The SAD pattern in Fig. 2(d) verifies the formation of the $\text{Pd}_{5+x}\text{ZnSe}$ phase and shows that it forms a specific orientation relationship with the substrate. Diffraction spots due to reflections from the $\text{ZnSe}[110]$ axis are marked by black dots, while the principal reflections from the $\text{Pd}_{5+x}\text{ZnSe}[110]$ zone axis are indicated by lines. The orientation relationship was found to be $\text{ZnSe}[110] \parallel \text{Pd}_{5+x}\text{ZnSe}[100]$ and $\text{ZnSe}(002) \parallel \text{Pd}_{5+x}\text{ZnSe}(001)$. The interplanar spacing of $\text{Pd}_{5+x}\text{ZnSe}(010)$ is nearly twice that of $\text{ZnSe}(220)$; therefore the lattice mismatch at the interface is only 1.3%. This small mismatch leads to a lower energy interface explaining the tendency for the phase to grow epitaxially.

RBS and XRD results show that the ternary $\text{Pd}_{5+x}\text{ZnSe}$ phase is stable up to 450 °C. Above 450 °C the Pd film begins to form islands on the surface and Pd diffuses completely through the ZnSe layer to react with the GaAs layer below. Figure 3 shows TEM micrographs of the microstructure after annealing at 500 °C for 30 min. The Pd remaining at the surface of the sample has formed a Pd_7Se_4 phase, as

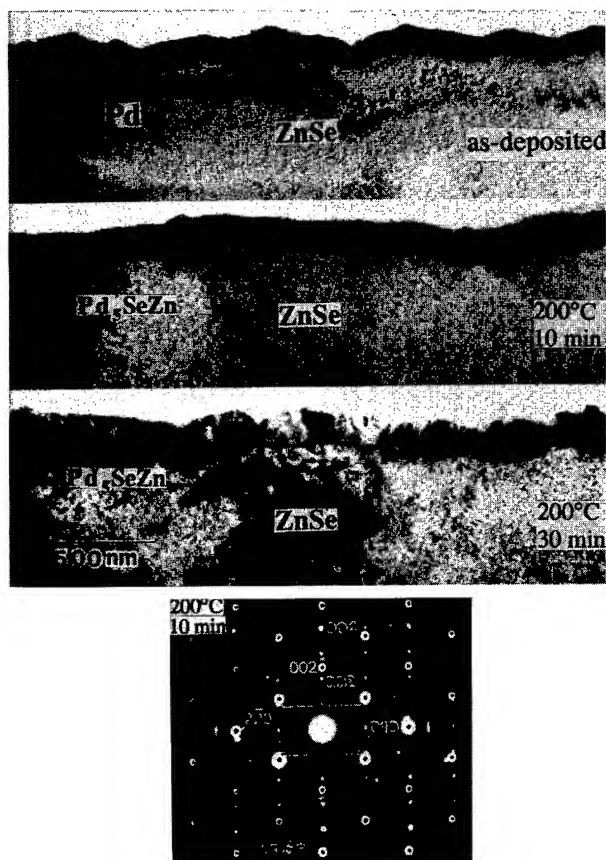


FIG. 2. TEM micrographs showing the progression of phase formation at the thin film Pd/ZnSe interface. (a) Morphology of the as-deposited sample, (b) $\text{Pd}_{5+x}\text{ZnSe}$ layer formed after annealing for 10 min at 200 °C, (c) increase in thickness of the $\text{Pd}_{5+x}\text{ZnSe}$ layer after annealing for 30 min at 200 °C, (d) selected area diffraction showing orientational relationship of the reacted $\text{Pd}_{5+x}\text{ZnSe}$ phase with the substrate. The black dots indicate the reflections from the ZnSe [110] axis. The lines denote the principal reflections of the $\text{Pd}_{5+x}\text{ZnSe}$ [100] zone axis.

identified by SAD and XRD. At the GaAs/ZnSe interface large precipitates which are a Pd–Ga phase have formed. In some locations large reacted areas, such as those shown in Fig. 3(b), are observed. These areas contain Pd, Zn, Se, Ga, and As, and are not a single phase, but a mixture of several phases comprised of these elements. The details of this interfacial reaction will be discussed in a future publication.

B. Pt/ZnSe

No significant reaction or interdiffusion is seen by RBS or XRD in the Pt/ZnSe system until it is annealed at temperatures of 550 °C or above. Annealing at 450 °C for 72 h results in agglomeration of the Pt on the ZnSe surface, but again there is no new phase formation detectable by RBS or XRD. Analysis by cross-sectional TEM indicates that there is a small amount of interdiffusion and some very small precipitates of a new phase after annealing at 450 °C. Figure 4 shows RBS spectra for Pt films annealed at 575 °C for 10 and 30 min compared to the as-deposited spectrum. After 10 min some Zn has diffused through the Pt layer and after annealing for 30 min the Zn concentration has increased.

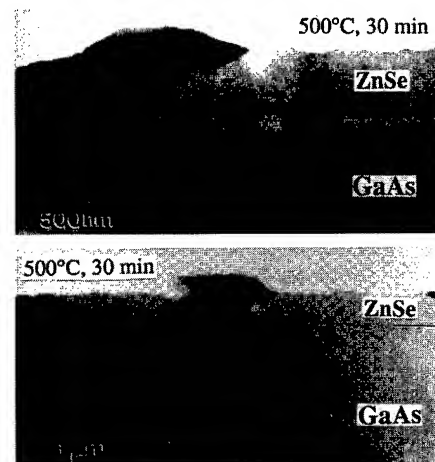


FIG. 3. TEM micrographs showing interface deterioration after annealing at 500 °C for 30 min. (a) Pd–Ga precipitates at the ZnSe/GaAs interface; (b) large reacted area at the ZnSe/GaAs interface.

Using RUMP, a RBS simulation program,¹⁴ the concentration of Zn in the Pt is estimated to be 20% after annealing for 30 min at 575 °C. Although there is a significant amount of Zn in the layer, no Zn–Pt compound can be identified by XRD. In addition, there is no indiffusion of Pt to the ZnSe/GaAs interface as there is in the case of Pd. The Pt diffraction peaks do shift to lower d values after annealing indicating a decrease in the lattice parameter of 0.017 Å. The formation of a solid solution of Pt:Zn with the smaller atoms substituting for the Pt can contribute to such a decrease in the lattice parameter. The Pt diffraction peaks also become narrower indicating grain growth in the film or a reduction of inhomogeneous strain in the layer. Most of the narrowing occurs in the first 10 min of annealing. SAD verifies that no Zn–Pt compounds form. After annealing for 30 min at 575 °C, RBS shows that the onset of the Pt peak has moved to a slightly lower energy. RUMP simulation shows that this is due to an approximately 50 Å layer of Zn that accumulates on the surface, thus showing that Zn is the main diffusing species during the reaction. Further evidence that Zn is diffusing to the surface is seen in the RBS spectrum of the sample an-

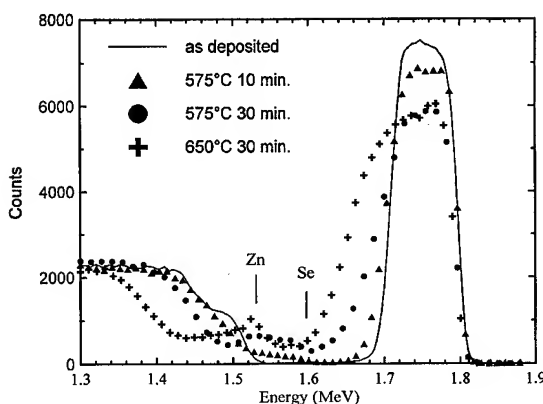


FIG. 4. RBS spectra of Pt on ZnSe annealed at 575 and 650 °C for various times.

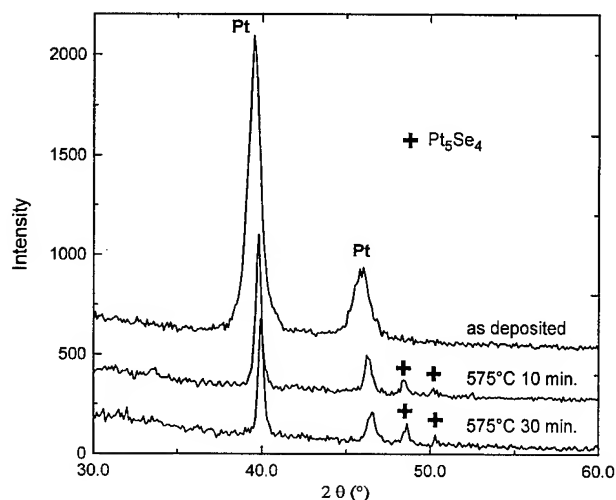


FIG. 5. X-ray spectra of Pt on ZnSe for the as-deposited sample and for samples annealed for various times at 575 °C.

nealed at 650 °C. In this spectrum we clearly see that the onset of the Pt peak occurs at a lower energy and there is large peak indicating the presence of Zn at the surface of the sample.

While no Pt–Zn phases are detected, a binary Pt–Se phase does form. XRD identifies the presence of the Pt_5Se_4 phase seen in Fig. 5. Pt_5Se_4 is a monoclinic phase with $a = 6.577 \text{ \AA}$, $b = 4.610 \text{ \AA}$, $c = 11.122 \text{ \AA}$, and $\beta = 101.59^\circ$. After annealing for 10 min, the (302) peak appears at $2\theta = 48.5^\circ$. After 30 min, the intensity of the (302) peak increases and the (123) peak at $2\theta = 50.5^\circ$ appears. There is also a decrease in the Pt peak heights, indicating that Pt was consumed in the reaction. The small intensity and different orientations of the Pt_5Se_4 peaks indicate that the phase is not epitaxially oriented to the substrate. The presence of this phase is verified by SAD as shown in Fig. 6(c). The particular grain chosen for SAD gives an orientational relationship with the substrate of $\text{Pt}_5\text{Se}_4[1\bar{2}0] \parallel \text{ZnSe}[110]$. The grains are randomly oriented with respect to the substrate. The micrograph in Fig. 6(b) shows that after annealing for 30 min the Pt_5Se_4 has formed a continuous layer, although the thickness varies between 33 and 73 nm. The reaction of the ZnSe with the Pt seems to affect the ZnSe layer differently than the Pd reaction. It is observed that the ion milling, which is required to form the TEM samples, results in some voids in the ZnSe near the interface, which is the light area directly below the Pt_5Se_4 layer in Fig. 6(b). This would suggest a compositional variation in the region that is related to the outdiffusion of Zn or that the reaction between Pt and Se has weakened the ZnSe layer in this region.

Annealing at higher temperatures results in increased Zn outdiffusion and the formation of Pt–Zn phases. Figure 7 shows an XRD spectrum from a Pt film annealed at 675 °C for 30 min. There is a number of Pt–Zn phases to which these peaks may belong, and the exact Pt–Zn phases present have not been identified. After this annealing cycle, the peaks for the Pt_5Se_4 layer are no longer evident indicating

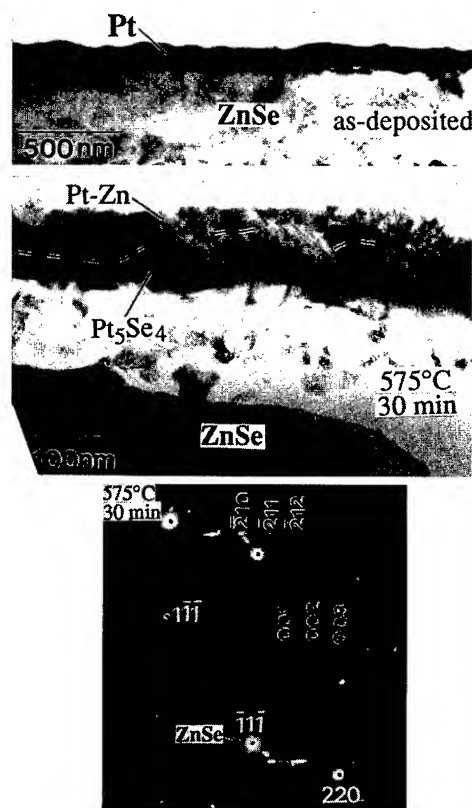


FIG. 6. TEM micrographs showing the progression of phase formation in the Pt/ZnSe system during anneals at 575 °C. (a) Morphology of as-deposited Pt on the ZnSe substrate, (b) Pt_5Se_4 layer formed after annealing at 575 °C for 30 min, (c) selected area diffraction showing orientational relationship of the Pt_5Se_4 grain with the substrate. $\text{Pt}_5\text{Se}_4[1\bar{2}0] \parallel \text{ZnSe}[110]$.

that this phase no longer exists at the interface. It was not possible to perform TEM for the phase identification because the metal layer repeatedly peeled off during sample preparation. After annealing at 575 °C the interface was weakened by the formation of voids. During annealing at the higher temperature the interface further deteriorated and TEM sample preparation was no longer possible.

The reaction of Pt on ZnSe is very sensitive to contaminants. Consistent reactions could not be obtained without a comprehensive cleaning. To get consistent behavior the samples must be solvent cleaned, etched in 49% HF for 30 s, and backspattered with Ar^+ ions for 6 min. The backspattering process removes approximately 350 Å of ZnSe. In addition, the Pt target was presputtered for 2 min before deposition. This treatment cleans contaminants off the surface and reduces any contamination in the Pt film deposited from the target. The observed nonuniformity in the thickness of the Pt_5Se_4 layer may be due to the presence of residual local contaminants on the ZnSe surface that slow down or inhibit the reaction.

C. Ni/ZnSe

The first observation of changes at the Ni/ZnSe interface by RBS occurs after annealing at 350 °C for 72 h with outdiffusion of Zn and Se. No changes could be detected by

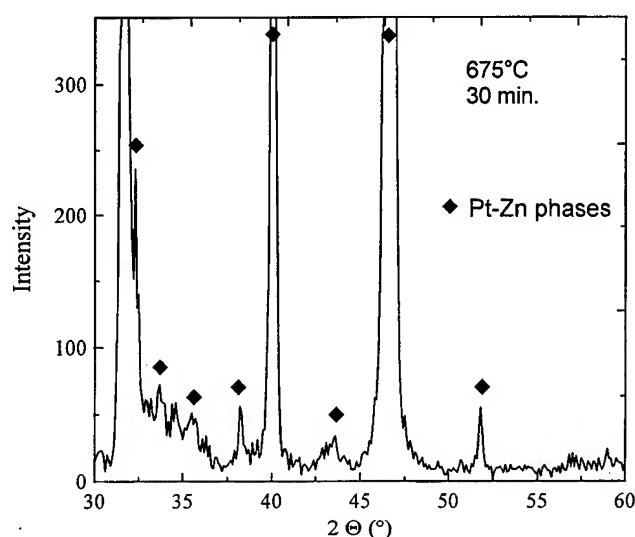


FIG. 7. X-ray spectra showing peaks of various Pt-Zn phases appearing in the Pt/ZnSe system after annealing at 675 °C for 30 min.

RBS or XRD after annealing for only 30 min at this temperature. TEM, however, provides an indication that there is formation of up to 50 Å diameter precipitates at the interface after 30 min of annealing at 300 °C.¹⁵ Figure 8 shows RBS spectra of Ni thin films annealed for 30 min at 400 and 425 °C compared to 1 h at 450 °C and an as-deposited Ni film. The initial Ni thickness was 700 Å. After annealing at 400 °C, Zn has diffused out to the surface and Se has diffused partially through the layer. Note that after annealing the onset of the Ni signal was pushed back to lower energy, indicating that Zn has accumulated on the surface of the film. Annealing at 425 °C results in increased interdiffusion and more Zn has accumulated at the surface. This can be observed because the onset of the Ni peak is at a lower energy. There is an increase in the amount of Se in the layer as well, although it has still not diffused completely through the Ni layer. This indicates that the reacted morphology is a layered structure.

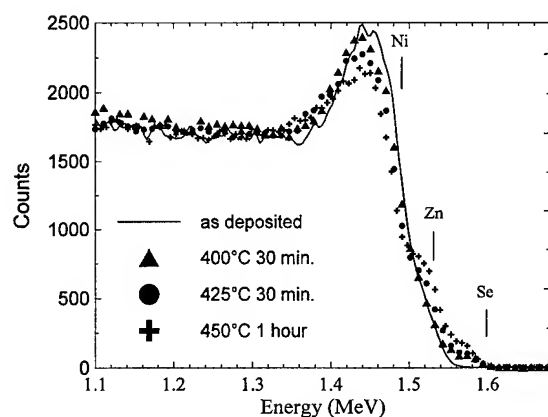


FIG. 8. RBS spectra of Ni on ZnSe annealed between 400 and 450 °C for various times.

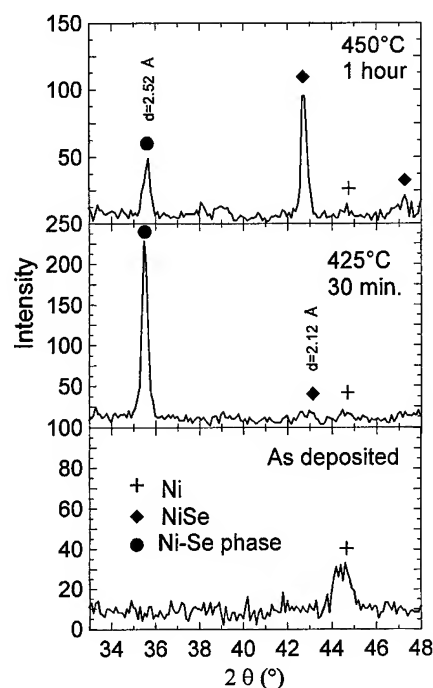


FIG. 9. X-ray spectra of Ni/ZnSe for the as-deposited sample and for samples annealed at 425 and 450 °C.

After annealing at 425 °C, new peaks are observed in XRD that correspond to a Ni-Se phase. Figure 9 shows XRD spectra from before and after annealing. The new peak at $2\theta=35.6^\circ$ corresponds to $d=2.52$ Å. The advent of this peak is coincident with that of a second peak at $2\theta=75.4^\circ$ ($d=1.26$ Å). EDS in the TEM shows that the ratio of Ni to Se in this phase ranges between 1 and 1.5. The XRD and SAD patterns are consistent but they cannot be correlated with any Ni-Se phase in the literature. This phase forms as separated grains, as shown in the TEM micrograph in Fig. 10. The grains are approximately 100 nm wide and extend approximately 50 nm into the ZnSe. In plan view, the diameter of the grains varies between 50 and 100 nm. This phase does not extend to the surface, which agrees with the RBS observation that Se has not diffused out to the surface. Between these grains and above them is a Ni-Zn solid solution.

Annealing at 450 °C results in the observation of two new XRD peaks at $d=2.11$ Å and $d=1.923$ Å as seen in Fig. 9. This corresponds to the formation of NiSe. The observed peaks are the (002) and (012) peaks, respectively. The intensity of these peaks increases after annealing at 450 °C for 1 h, while the peak for the Ni_xSe ($1 < x < 1.5$) phase decreases. EDS of the TEM sample shows that the ratio of Ni to Se in this new phase ranges between 0.6 and 1. This second NiSe phase was observed previously in a Ni-Se thin film,¹⁶ and is cubic with $a=4.3$ Å. Peaks for both phases are observed in samples annealed at 350 °C for 72 h with approximately equal intensities. The Ni_xSe phase forms initially during the reaction. With longer annealing or higher temperatures the phase transforms to NiSe as more Se is incorporated into the grains. This indicates that the Ni_xSe phase may be metastable.

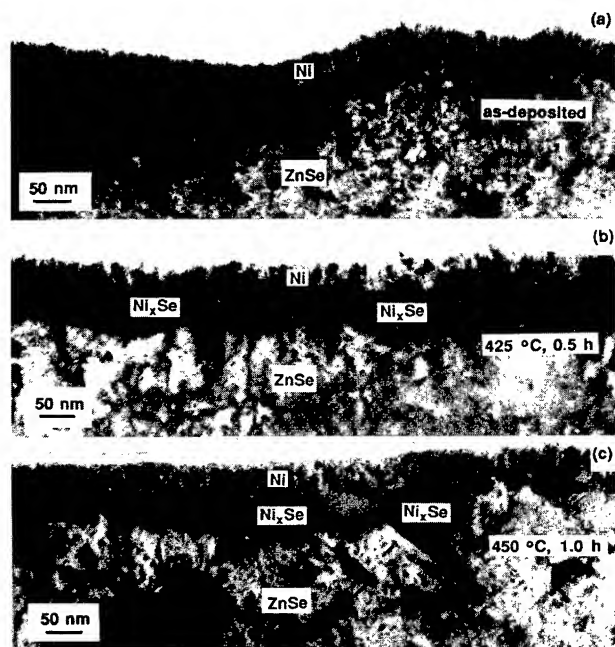


FIG. 10. TEM micrographs showing the progression of phase formation in the Ni/ZnSe system during anneals at 425 and 450 °C. (a) As-deposited Ni on ZnSe, (b) grains of the Ni-Se phase formed at the interface after annealing at 425 °C for 30 min, (c) grains of NiSe formed at the interface after annealing at 450 °C for 1 h.

Figure 10(c) shows that the NiSe phase also forms as separated grains. These grains are larger than those formed at 425 °C, and have diameters in plan view between 100 and 200 nm. Above this layer of separated grains and between them there is still a solid solution of Zn in Ni and within this region are small inclusions of NiSe. The thin layer of Zn at the surface observed by RBS could not be found by TEM due to the high vapor pressure of Zn and the ion milling required for sample preparation. Careful TEM observation of a plan view specimen annealed at 425 °C for 30 min shows that the NiSe phase is already present in small amounts at the lower temperature. A more detailed discussion of the microstructure can be found in Ref. 15.

IV. DISCUSSION

The lowest initial reaction temperature of the near-noble metals on ZnSe occurs for Pd which reacts at 200 °C. The highest is found for Pt, which reacts at 575 °C. This trend is also observed for Si and GaAs, as shown in Fig. 11. Figure 11 compares the heat of vaporization of the metal to the initial reaction temperature for Si, GaAs, and ZnSe. The reaction temperature increases with the heat of vaporization of the metal, which is a measure of the binding energy between the metal atoms. This implies that more thermal energy is needed to begin the reaction when the metal is more strongly bonded. The reaction temperature also increases with the heat of formation of the semiconductor. Si has the lowest reaction temperatures and the lowest heat of formation. There are also trends regarding orientation, sensitivity to

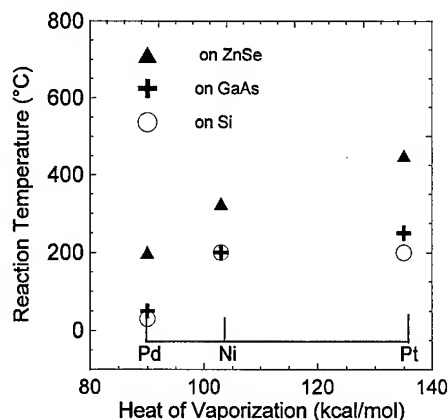


FIG. 11. Initial reaction temperature vs metal heat of vaporization for reactions of the near-noble metals on Si, GaAs, and ZnSe.

contaminants, and morphology in the reactions of the near-noble metals on Si, GaAs, and ZnSe. These trends will be further discussed below.

A. Palladium

The reactions of Pd are characterized by oriented films that often form in a granular structure. Pd_2Si forms with an epitaxial relationship to a $\langle 111 \rangle$ Si substrate while the ternary compound on GaAs tends to form in slightly misoriented grains.^{17,18} $\text{Pd}_{5+x}\text{ZnSe}$ also forms in oriented grains that become more epitaxial as the reaction proceeds. Also, as on GaAs, Pd on ZnSe forms an oriented ternary compound. In the GaAs system, a second ternary phase forms when it is annealed at higher temperatures. This is then followed by the appearance of M_xGa_y and M_xAs_y binary phases. In the Pd/ZnSe system, no second ternary phase is observed. Above 450 °C the reaction is no longer controllable and Pd diffuses to the ZnSe/GaAs interface to form Pd-Ga precipitates in a nonplanar reaction. Reactions of Pd metal on Si, GaAs, and ZnSe are not affected by impurities or surface cleanliness.

The layer of $\text{Pd}_{5+x}\text{ZnSe}$ grows very quickly in the first 10 min of reaction and then the growth rate slows. This is typical of a diffusion controlled reaction, which is characterized by a growth rate of the reacted layer that is proportional to the square root of time. The proportionality factor is the interdiffusion coefficient which is exponentially dependent on the activation enthalpy of the reaction. This activation enthalpy can be determined from the change in reacted layer thickness during a series of isothermal anneals. It is difficult to determine the exact layer thickness from RBS because of the compositional variation in the reacted layer, but the activation enthalpy for the reaction could be determined from a more detailed TEM study.

The initial Pd reaction, which occurs at 200 °C, may be useful for electrical contacts. The reaction takes place at a low enough temperature that the electrical characteristics of the ZnSe film should not be affected. The reaction provides a planar, epitaxial interface that is desired for electrical contacts. In addition, the reaction is not very sensitive to contaminants or oxides and therefore the surface does not re-

quire special preparation, such as backspattering or other ultrahigh vacuum cleaning. In order to develop Pd-based contacts on ZnSe detailed electrical measurements are needed to determine the Schottky barrier height or contact resistivity of the ternary phase on *n*- and *p*-type ZnSe. The formation of this phase is very similar to the initial formation of the ternary phase on GaAs. This ternary phase formation on GaAs has been utilized to form ohmic contacts using solid phase epitaxial regrowth. This process may also be useful for ohmic contacts to *p*-type ZnSe using Li, for example, as a dopant.

B. Platinum

The reaction of Pt on both ZnSe and GaAs is characterized by the formation of a metal-anion phase at the metal/semiconductor interface and outdiffusion of the cation.¹⁹ In the GaAs case the outdiffusion of Ga also results in the formation of Pt-Ga phases, but no metal-cation phases form in the ZnSe case. This may be due to the high solubility of Zn in Pt. In both cases the main moving species appears to be the cation and the Pt. The cation diffuses out, whereas the Pt diffuses in and reacts with the anion which remains at the interface. This vertical phase stratification is expected in films that have thicknesses greater than the grain size.¹⁰ It is likely that the reaction will result in lateral phase separation if the thickness of the Pt layer is less than the grain size of the reacted phase.

The Pt reaction is very sensitive to contamination or oxides at the interface between the ZnSe and Pt. Consistent reaction behavior is obtained only if the samples are HF etched and Ar⁺ backspattered prior to deposition of the metal. A HF dip etch alone does not remove enough of the C and O from the surface of the ZnSe. X-ray photoelectron spectroscopy measurements have shown that sputtering removes more C and O from the surface. The variation in thickness of the Pt₅Se₄ layer is an indication that, even after sputtering the surface, the reaction does not proceed at a uniform rate across the sample. This may be due to areas of contamination that remain on the surface. The reaction of Pt on ZnSe appears to be more sensitive to contamination than the reactions on Si or GaAs. The reaction of Pt on Si is sensitive to impurities in the Pt film.²⁰ While the presence of oxygen in the Pt film was observed to change the temporal formation sequence of observed silicide phases, the reaction on GaAs is not sensitive to native oxides,²¹ but it is slowed by the presence of thick oxides.²²

The reaction of Pt on ZnSe is not amenable to contact formation by alloying. The temperature for reaction is too high and will adversely affect the electrical properties of the ZnSe. In addition, although the reacted Pt₅Se₄ layer is complete, the thickness is not uniform and will lead to inconsistencies in the electrical properties.

C. Nickel

In the reaction on ZnSe, no ternary phase is observed; rather, separated grains of a Ni-Se binary phase begin to form initially. This is in contrast to Ni on GaAs which begins

with the formation of a ternary phase Ni₃GaAs, followed by the formation of laterally separated grains of binary Ni-Ga and Ni-As phases at higher temperatures. The formation of NiAs grains is observed at a slightly higher temperature in the Ni/GaAs reaction and these are also separated, but by a Ni-Ga binary phase. The reaction of Ni on ZnSe is not extremely sensitive to the surface preparation, as Ar⁺ backspattering is not necessary for the new phases to form. The reactions of Ni on Si and GaAs are also not sensitive to oxides unless they are thick.

This reaction is also not desired for electrical contacts. Again the reaction temperature is too high to maintain the electrical characteristics of the ZnSe. Also, the reaction does not provide a laterally homogeneous film for a contact. The laterally separated grains that are formed will provide contacts parallel to areas where no phase has formed, most likely leading to inconsistent and unpredictable results.

V. CONCLUSIONS

The initial reactions of Pd and Pt on ZnSe occur at 200 and 575 °C, respectively, and are very similar to the reactions of Pd and Pt on GaAs. Pd tends to first form a ternary oriented phase, Pd_{5+x}ZnSe, while Pt reacts to form a Pt₅Se₄ layer at the interface. The Pd reaction appears to be diffusion controlled, since the growth of the reacted layer slows with time. The reaction of Pt is very sensitive to contaminants and oxides at the interface (as are its reactions on Si and GaAs), whereas the Pd reaction is uninhibited by small amounts of contamination or oxide. The Ni reaction on ZnSe, however, is different from that on GaAs. Ni begins to react at 300–350 °C to form separated grains of a Ni-Se phase at the interface between the Ni and ZnSe. This is in contrast to the reaction of Ni on GaAs which initially forms a ternary Ni₃GaAs phase. The increase in reaction temperature from Pd to Ni to Pt parallels the increase in binding energy of the metal. In addition, the reactions on ZnSe take place at higher temperatures than the corresponding reactions on Si and GaAs. This trend follows an increase in the semiconductor heat of formation from Si to GaAs to ZnSe.

The reaction that provides an interesting possibility for the formation of electrical contacts by alloying is that of Pd on ZnSe. The Ni and Pt reactions take place at a temperature that is above the temperature at which a compensating defect forms and degrades the electrical characteristics of ZnSe; in addition these reactions do not result in a laterally homogeneous contact. The Pd reaction is particularly interesting because it is very planar and controllable and the ternary phase forms with an epitaxial relationship with the substrate. This type of behavior is similar to the Pd ternary formation on GaAs and may lead to the development of various Pd-based alloyed ohmic or Schottky contact structures similar to those developed for GaAs. However, the electrical characteristics of the contact between Pd_{5+x}ZnSe and ZnSe are not yet determined. This phase may prove to be a good ohmic or Schottky contact to ZnSe. The stability of Pt and Ni above temperatures at which ZnSe devices will be operated makes them suitable candidates for stable Schottky contacts.

ACKNOWLEDGMENTS

The authors wish to thank W. Swider for TEM sample preparation and J. M. Walker for his assistance with MOCVD growth. One author (K.J.D.) was supported by an AT&T Ph.D. fellowship. This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Science Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098. The use of the facilities of the National Center for Electron Microscopy is also acknowledged.

- ¹R. L. Gunshor, J. Han, A. V. Nurmikko, and A. Salokatve, *J. Cryst. Growth* **150**, 790 (1995).
- ²J. Qiu, H. Cheng, J. M. DePuydt, and M. A. Haase, *J. Cryst. Growth* **127**, 279 (1993).
- ³C. Boney, Z. Yu, J. W. H. Rowland, W. C. Hughes, J. J. W. Cook, and J. F. Schetzina, *J. Vac. Sci. Technol. B* **14**, 2259 (1996).
- ⁴K. Ohkawa, T. Karasawa, and T. Mitsuyu, *Jpn. J. Appl. Phys.* **1** **30**, L152 (1991).
- ⁵L. C. Wang, P. H. Hao, and B. J. Wu, *Appl. Phys. Lett.* **67**, 509 (1995).
- ⁶M. Heuken, J. Sollner, F. E. G. Guimaraes, K. Marquardt, and K. Heime, *Appl. Phys. Lett.* **60**, 1694 (1992).
- ⁷A. L. Chen, W. Walukiewicz, K. Duxstad, and E. E. Haller, *Appl. Phys. Lett.* **68**, 1522 (1996).
- ⁸P. S. Ho, T. Y. Tan, J. E. Lewis, and G. W. Rubloff, *J. Vac. Sci. Technol.* **16**, 1120 (1979).
- ⁹C. Canali, F. Catellani, G. Ottaviani, and M. Prudenziati, *Appl. Phys. Lett.* **33**, 187 (1978).
- ¹⁰T. Sands, V. G. Keramidas, K. M. Yu, J. Washburn, and K. Krishnan, *J. Appl. Phys.* **62**, 2070 (1987).
- ¹¹K. J. Duxstad, E. E. Haller, K. M. Yu, E. D. Bourret, J. M. Walker, X. W. Lin, and J. Washburn, *Appl. Phys. Lett.* **67**, 947 (1995).
- ¹²E. D. Bourret, F. X. Zach, K. M. Yu, and J. M. Walker, *J. Cryst. Growth* **147**, 47 (1995).
- ¹³V. M. El-Boragy and K. Schubert, *Z. Metallkd.* **61**, 579 (1970).
- ¹⁴L. R. Doolittle, *Nucl. Instrum. Methods Phys. Res. B* **15**, 227 (1986).
- ¹⁵Sergei Ruvimov, Zuzanna Liliental-Weber, Wendy Swider, Jack Washburn, Edith Bourret, Kristin J. Duxstad, and Eugene E. Haller, *J. Vac. Sci. Technol. B* **15**, 247 (1997).
- ¹⁶A. K. Singh, M. M. Srivastava, and O. N. Srivastava, *J. Less-Common Met.* **57**, 225 (1976).
- ¹⁷D. Sigurd, R. W. Bower, W. F. VanDerWeg, and J. W. Mayer, *Thin Solid Films* **19**, 319 (1974).
- ¹⁸T. S. Kuan, J. L. Freeouf, P. E. Batson, and E. L. Wilkie, *J. Appl. Phys.* **58**, 1519 (1985).
- ¹⁹C. Fontaine, T. Okumura, and K. N. Tu, *J. Appl. Phys.* **53**, 1404 (1983).
- ²⁰C. A. Crider and J. M. Poate, *Appl. Phys. Lett.* **36**, 417 (1980).
- ²¹T. Sands, V. G. Keramidas, A. J. Yu, K.-M. Yu, R. Gronsky, and J. Washburn, *J. Mater. Res.* **2**, 262 (1987).
- ²²V. Kumar, *J. Phys. Chem. Solids* **36**, 535 (1975).

Ultrathin cobalt silicide layers formed by rapid thermal processing of metal on amorphous silicon

A. H. M. Kamal, M. J. Rack, M. N. Kozicki,^{a)} and D. K. Ferry

Center for Solid State Electronics Research, Arizona State University, Tempe, Arizona 85287-6206

J. Lützen

Science and Engineering of Materials Program, Arizona State University, Tempe, Arizona 85287-1704

J. A. Hallmark

Phoenix Corporate Research Laboratories, Motorola Inc., Tempe, Arizona 85284

(Received 25 July 1996; accepted 9 May 1997)

The formation of thin (<15 nm) cobalt silicide layers on amorphous silicon (a -Si) has been investigated as a part of a program involving the fabrication of poly-Si-based single electron transistors. Rapid thermal annealing was used to promote the reaction between the a -Si and cobalt metal films. The effects of three most important factors on silicide formation during annealing (temperature, temperature ramp rate, and anneal time) were studied using both electrical (sheet resistance) and surface analysis (Auger electron spectroscopy) techniques. In particular, the impact of temperature ramp rate on silicide formation has been addressed. The optimization of the thin silicide process has been achieved by a statistical analysis. Transmission electron microscopy cross-sectional analysis has been performed to determine the thickness of the film. © 1997 American Vacuum Society. [S0734-211X(97)12604-X]

I. INTRODUCTION

The small dimensions of devices used in very large scale integration necessitate the application of refractory metal silicides in source-drain and gate-interconnection areas to reduce the contact resistance and Ohmic voltage drop, respectively. The self-aligned silicide process (SALICIDE) has been reported as an effective solution for minimizing gate and source/drain resistance with subsequent increase in circuit speed.¹⁻³ Among the possible silicide candidates suitable for the SALICIDE process, TiSi_2 and CoSi_2 have been identified as the most promising from the view point of resistivity: $14\text{--}20\ \mu\Omega\text{ cm}$ for TiSi_2 and $16\text{--}20\ \mu\Omega\text{ cm}$ for CoSi_2 .⁴ The main disadvantage of TiSi_2 is the lateral growth, which occurs during a SALICIDE process and can cause gate-to-source bridging.⁵ In contrast, a minimum silicide lateral growth is expected during CoSi_2 formation because cobalt (Co) is the dominant moving species in Co_2Si and CoSi_2 formation.⁶ Unlike titanium, cobalt does not reduce SiO_2 during a SALICIDE process.

Since most metals are liable to form metal oxides before they are transformed into silicides during annealing at high temperatures, conventional furnace annealing has been replaced by rapid thermal annealing (RTA), at the expense of throughput and cost, to achieve precise control over the annealing ambient.⁵ Freitas and Swart⁴ studied the influence of the impurity content in Co films and the presence of native oxide at the Co/Si interface on cobalt silicide formation and concluded that for Co films with low contamination levels, even thick native oxide can be dissolved, and good quality CoSi_2 is formed. Whereas for Co films with a high impurity level, the silicidation, in the presence of thick oxide, can be

performed, delayed, or even inhibited, depending on film thickness and annealing conditions.

Information extracted from studies made on either thick poly-Si or single-crystal substrates was found to be inadequate when the silicide formation of thin poly-Si (≤ 40 nm) was attempted. Thin poly-Si silicide films, however, can be used in the fabrication of poly-Si-based single electron transistors (SET) first proposed by Yano *et al.*,⁷ to reduce the operating voltage of the device. The operating voltage of the original SET was about 60–90 V due to the use of the high resistivity poly-Si source, drain, and thicker (150 nm) gate oxide. Replacement of this poly-Si layer by thin poly-Si silicide will drastically reduce the operating voltage because the series resistance of these single electron transistors will be reduced by two orders of magnitude. A rigorous study of silicide formation on the thin a -Si layer was conducted with an ultimate goal of fabricating a low-voltage poly-Si grain-based SET. In addition to being utilized in SET fabrication, thin cobalt silicide lines may emerge as an essential wiring material for next-generation nanoscale devices. The lattice constant of CoSi_2 is 0.536 nm ,⁸ which is almost equal to that of Si, 0.543 nm . This close matching of these two lattices results in very flat CoSi_2 lines.⁸ The improved planarity offered by thin cobalt silicides will make them suitable for next-generation nanoscale device fabrication.

In this study, cobalt polycide films were formed on as-deposited amorphous-Si layers because it has been reported⁹ recently that polycide gates formed on as-deposited amorphous silicon exhibit greater thermal stability than those formed on as-deposited polycrystalline silicon. Moreover, the source, drain, and gate of the SET structure, mentioned before, are constructed using thin (10–20 nm) poly-Si films. Poly-Si of this structure is deposited as amorphous Si and

^{a)}Electronic mail: michael.kozicki@asu.edu

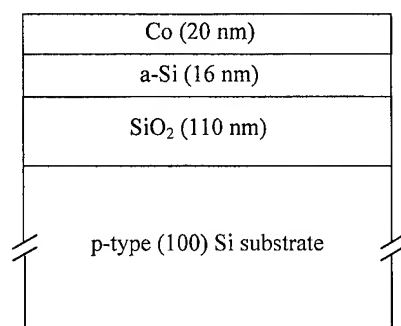


FIG. 1. Different films grown or deposited on Si substrates to prepare all samples.

then recrystallized to achieve uniformity and controllability in film thickness.

II. EXPERIMENT

Figure 1 shows the layers grown or deposited on two 4 in., *p*-type, (100) oriented Si substrates. After standard RCA cleaning, thermal oxides, ~110 nm thick, were grown on both substrates. Amorphous Si (*a*-Si) was deposited by low-pressure chemical vapor deposition (LPCVD) at 540 °C and 250 mTorr with a silane flow rate of 100 sccm. The thickness of the amorphous films was measured by spectroscopic ellipsometry. All samples were dipped in HF:H₂O = 1:100 for 120 s before loading them in the e-beam evaporator to remove native oxide. Cobalt films were deposited with a base pressure of 3×10^{-6} Torr and a deposition rate of 0.03 nm/s. A low deposition rate was maintained to achieve uniformity of deposition.

Rapid thermal annealing was carried out in a Tamarack model 180-M manual radiant impulse processor using N₂ as the ambient gas. In this RTA system, 15 tungsten-halide lamps are staggered in two rows in the upper section of the chamber. The sample is placed in the lower section. Unreacted Co was etched in HCl:H₂O₂:H₂O = 3:1:3 solution at about 80 °C, which does not etch cobalt silicide.¹⁰ Sheet resistances were measured by a VEECO FPP-5000 automatic four-point probe. The silicide phase was characterized by a Perkin-Elmer PHI 600 Auger electron spectroscopy. Auger electron spectroscopy (AES) was also used to determine the composition of an as-deposited Co/*a*-Si sample.

32 samples were prepared for a 2³ factorial analysis. Although the same deposition parameters were used during LPCVD of *a*-Si on 110 nm of oxide, *a*-Si was deposited in two different runs and, therefore, all treatment combinations of the 2³ factorial design were run in two blocks with highest-order interactions confounded with the blocks. Thus, differences in sheet resistances, which may arise due to the use of two different LPCVD runs, have been distinguished as a block effect. To obtain an estimate of error, four replicates of each treatment combination were run. A 2³ factorial design¹¹ having four replicates and highest-order interaction confounded with blocks has been performed. Two values of

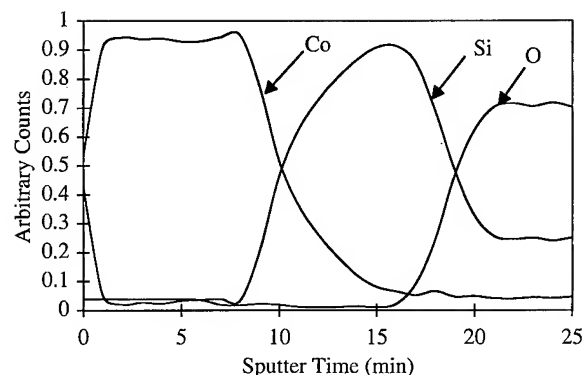


FIG. 2. AES depth profile for the as-deposited sample with 40 nm cobalt deposited on 40 nm of amorphous silicon.

temperature, temperature ramp rate, and anneal time during RTA were chosen as 650 and 750 °C, about 130 and 145 °C/s, and 10 and 30 s, respectively.

Transmission electron microscopy (TEM) cross-sectional analysis of one of the samples has been performed to determine the thickness of the cobalt silicide layer and the uniformity of the CoSi₂/SiO₂ interface.

III. RESULTS AND DISCUSSIONS

As a preliminary study, silicide formation was attempted with slower ramp rates (<100 °C/s) with or without preheating at 450 °C for 5–10 s. In one of such attempts, 3 nm Co was deposited on 8 nm *a*-Si. Rapid thermal annealing at 750 °C of this Co/*a*-Si system did not result in silicide formation. The thin layer of Co reacted with ambient oxygen to transform into Co oxide. It was thought that a thicker Co film might prevent ambient oxygen from oxidizing the Co species and improve the mechanism of silicide formation. Although the Co species located closer to the surface will have oxygen to react with, the Co species near the Co/Si interface will be free of oxygen. Co, from this oxygen-free region, will diffuse into the *a*-Si and form cobalt silicide during the annealing period. Therefore, a thicker Co film, 40 nm, was deposited over a 40 nm *a*-Si film. However, no silicide was achieved with slower ramp rates with or without preheating. The slower ramp rate might have transformed Co into cobalt oxide. It appeared that Co oxide, if formed anytime during annealing, cannot be reduced to silicide at an elevated temperature. These results for a thin *a*-Si film support the conclusion made by Yang and Chen¹⁰ and are contrary to a previous report,¹² which proposed that CoO on silicon can be reduced to silicide at temperatures above 750 °C. Figure 2 is the AES depth profile for the as-deposited sample, where 40 nm of Co was deposited on 40 nm of *a*-Si. The lower concentration of oxygen in this as-deposited film indicates that the oxygen incorporated in the Co film was not enough to affect the annealing process; it is ambient oxygen, which inhibits the silicide formation.

To overcome the low-temperature oxidation of Co, the influence of ramp rate during rapid thermal annealing on the silicide formation of thin poly-Si films was studied further.

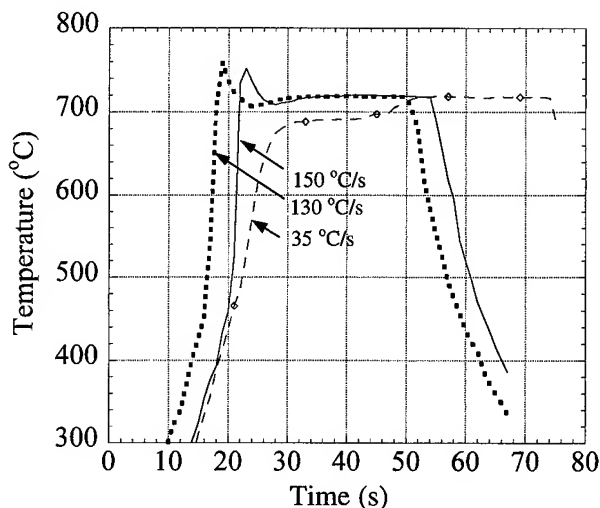


FIG. 3. Actual temperature rise with time during RTA for three different temperature ramp rates.

Figure 3 shows temperature rise with time of the RTA for three different selections of temperature ramp rate. SensArays's Series 1501 A-X-1001 Instrumented Wafers were used for temperature measurements. In these wafers, a type-K (chromel–alumel) thermocouple (TC) is bonded into the wafer surface, which allows direct measurement of the wafer temperature. Using these TC calibration wafers, the calibration stability of the pyrometer is assured by a wafer's temperature repeatability specification of $\pm 0.2^\circ\text{C}$. To the best of our knowledge, the effect of ramp rate on silicide formation has not been reported previously. A series of experiments, therefore, were performed to determine the effect of this factor. When the annealing temperature was reached with a ramp rate of $\sim 150^\circ\text{C/s}$, the sheet resistance measurements of these annealed films, the phase of which was unknown before the Auger analysis, dropped considerably to yield a resistivity about $36\ \mu\Omega\text{ cm}$. This improvement may be due to the fact that at higher ramp rate the reaction for silicide formation is greater than that of Co oxide formation. Figure 4 is the Auger analyzed depth profile for the sample that was annealed at 750°C for 30 s with a ramp rate of 130°C/s . From this profile, the formation of cobalt silicide is obvious;

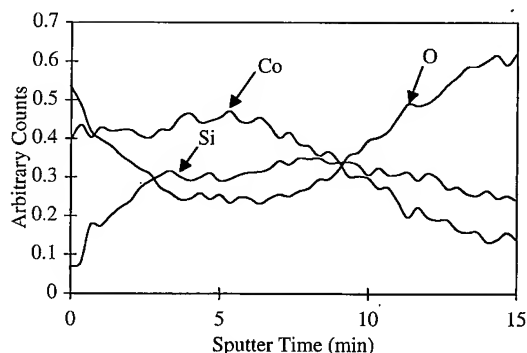


FIG. 4. AES depth profile for the sample that was annealed at 700°C for 30 s with a ramp rate of $\sim 130^\circ\text{C/s}$.

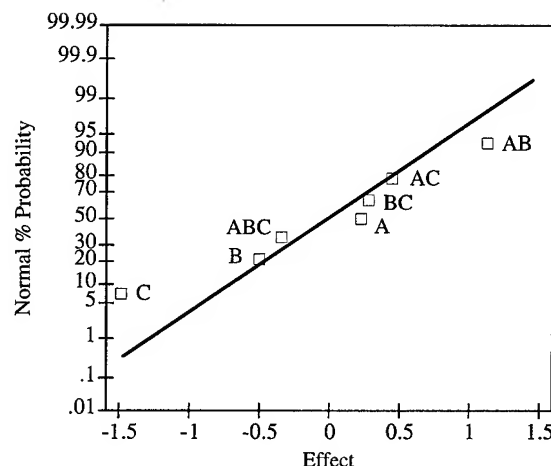


FIG. 5. Normal probability plot of effects for all treatment combinations.

however, due to oxygen incorporation, it is difficult to identify any of the phases of cobalt silicide, i.e., Co_2Si , CoSi , or CoSi_2 , clearly. The fact that oxygen from the ambient gets incorporated during annealing is evident from Fig. 4. Oxygen incorporation can be prevented by using a thin ($\sim 5\text{ nm}$) amorphous-Si capping layer on cobalt.^{13,14}

A. Model for estimating sheet resistance

The effects of process conditions on the formation of thin poly-Si silicide films were studied by 2^3 factorial design with temperature, temperature ramp rate, and anneal time as possible dominant factors. Since the purpose of the thin Co polycide is to reduce the source and drain resistance of a thin poly-Si film transistor, the sheet resistance was chosen as the response variable. Figure 5 is the normal probability plot of mean sheet resistance for all treatment combinations. A, B, and C, in Fig. 5, represent treatment combinations where temperature (A), temperature ramp rate (B), and anneal time (C) are set at their highest value. AB is the treatment combination for which temperature and temperature ramp rate are at their high level but anneal time is at its low level. It is obvious from Fig. 5 that the variations of sheet resistances are statistically significant only for the main effect C and interaction AB. These two effects, C and AB, have nonzero means and do not lie along the straight line. Average sheet resistances for all the other six treatment combinations fall on a straight line implying that these effects are negligible and their responses are normally distributed with mean-zero and variance σ^2 . The analysis of variance is detailed in Table I, where the sum of the square for the model is defined as

$$SS_{\text{Model}} = SS_A + SS_B + SS_C + SS_{AB} + SS_{BC} + SS_{AC} + SS_{ABC} \quad (1)$$

The value of R^2 (Ref. 11) is 72, which suggests that about 72% of the variability in the sheet resistance is explained by the model. The unexplained variability could be due to fac-

TABLE I. Analysis of variance for the 2^3 factorial design with ABC confounded with blocks.

Source	Sum of squares	Degrees of freedom	Mean square	F value	Prob. > F
Block	0.002 81	1	0.0028		
Model	33.33	7	4.76	8.4	<0.0001
Error	13.04	23	0.567		
Total	46.371	31			

$$R^2 = \frac{SS_{\text{Model}}}{SS_{\text{Total}}} = \frac{33.33}{46.37} = 0.72$$

tors such as randomness of the oxygen content in the N_2 ambient and thickness of the native oxide. Residuals, in this model, are given as

$$e = \rho_s - \hat{\rho}_s, \quad (2)$$

where ρ_s is the observed sheet resistance and $\hat{\rho}_s$ is the fitted sheet resistance.¹¹ Though the variation in sheet resistance, for some treatment combinations, is statistically significant, it is too small to have reasonable practical importance. We can, therefore, conclude that similar thin silicide films can be obtained by RTA with temperature, temperature ramp rate, and anneal time chosen in the range used for all the samples.

Figure 6 is the TEM cross section of a sample, the AES profile of which is shown in Fig. 4. According to this picture, the thickness of the cobalt silicide formed on the oxide is about 10 nm. The average sheet resistance of this film is about 36 Ω/sq , which yields a resistivity of about 36 $\mu\Omega \text{ cm}$. This resistivity is twice that of the standard cobalt disilicide. The increase in resistivity is surely due to the presence of oxygen in the film. The silicide/oxide interface is nonuniform, which confirms the existence of spatial variation in

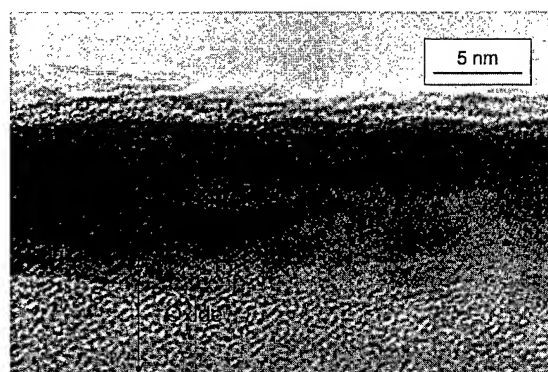


FIG. 6. TEM cross section of cobalt silicide formed on the $a\text{-Si}/\text{SiO}_2$ (16/100 nm) system.

silicide thickness in this film. The 10 nm cobalt silicide films are found to be thermally stable up to 850 $^{\circ}\text{C}$.

IV. CONCLUSION

The formation of a thin polycide film on a thin layer of $a\text{-Si}$ is governed by a mechanism that differs from the mechanism of forming silicide on thick poly-Si ($\geq 300 \text{ nm}$) or single-crystal Si. Preheating or a slower temperature ramp rate ($< 100 \text{ }^{\circ}\text{C/s}$) during annealing in a rapid thermal annealing system allows Co to react with ambient oxygen to produce Co oxide. Once Co oxide is formed, it cannot be reduced to silicide by annealing at 700–800 $^{\circ}\text{C}$. Silicidation is possible with higher temperature ramp rate ($\sim 150 \text{ }^{\circ}\text{C/s}$) even in an oxygen contaminated ambient. Identification of the silicide phases is difficult because of the presence of oxygen in our 10 nm silicide films. The thickness of the $a\text{-Si}$ layer ($\sim 16 \text{ nm}$) may not be sufficient for the formation of the CoSi_2 phase.

A 2^3 factorial design, with temperature, temperature ramp rate, and anneal time considered as the three factors of RTA, does not show significant variation in sheet resistances among the samples prepared by different treatment combinations. About 72% of the variability in the sheet resistance is explained by the model. The thickness of the silicide, determined by TEM analysis, is about 10 nm.

ACKNOWLEDGMENTS

The authors thank Dr. Max Sidorov of the Center for Solid State Science at ASU for the TEM cross-sectional analysis. This research is supported by the Phoenix Corporate Research Laboratory, Motorola Inc.

¹C. Y. Ting, Tech. Dig. Int. Electron Devices Meet. **84**, 110 (1984).

²S. P. Murarka, J. Vac. Sci. Technol. B **4**, 1325 (1986).

³S. P. Murarka, J. Vac. Sci. Technol. **17**, 775 (1980).

⁴W. J. Freitas and J. W. Swart, J. Electrochem. Soc. **138**, 3067 (1991).

⁵A. E. Morgan, E. K. Broadbent, M. Delfino, B. Coulman, and D. K. Sanda, J. Electrochem. Soc. **134**, 925 (1987).

⁶S. S. Lau, J. W. Mayer, and K. N. Tu, J. Appl. Phys. **49**, 4005 (1978).

⁷K. Yano, T. Ishii, T. Hashimoto, T. Kobayashi, F. Murai, and K. Seki, IEEE Trans. Electron Devices **41**, 1628 (1994).

⁸K. H. Robrock, K. N. Tu, D. W. Abraham, and J. B. Clabes, Appl. Phys. Lett. **54**, 1543 (1989).

⁹W. M. Chen, J. Lin, S. K. Banarjee, and J. C. Lee, J. Appl. Phys. **73**, 4712 (1993).

¹⁰Fann-Mei Yang and Mao-Chieh Chen, Thin Solid Films **207**, 75 (1992).

¹¹D. C. Montgomery, *Design and Analysis of Experiments* (Wiley, New York, 1991), Chaps. 7–10.

¹²J. Chevallier and A. N. Larsen, Appl. Phys. A **39**, 141 (1986).

¹³Y. S. Lou, C. Y. Wu, and H. C. Cheng, Solid-State Electron. **36**, 75 (1993).

¹⁴A. H. M. Kamal, Ph.D. thesis, Arizona State University (1997).

Electrical characterization of iridium Schottky contacts to silicon: Early stages of silicidation

F. J. Jiménez-Leube,^{a)} M. Clement, J. Sanz Maudes, and T. Rodríguez

Departamento de Tecnología Electrónica, ETSI Telecomunicación, UPM, Avenida Complutense s/n, 28040 Madrid, Spain

(Received 20 November 1996; accepted 4 April 1997)

High-quality Ir/*n*-Si Schottky diodes have been prepared. The diodes exhibit ideal behavior at room temperature according to thermionic emission theory with linear $\log(I) - V$ characteristics over eight current decades. Ideality factors less than 1.02 were obtained. Schottky barrier heights determined by forward $I - V$ and $C - V$ characterization are close to 0.9 eV. Current-voltage characteristics show deviations from the simple thermionic emission law at low temperature. This anomalous behavior is explained in terms of spatial inhomogeneities in the barrier height that result from the presence of a polycrystalline interlayer at the metal-semiconductor interface. The flat-band Schottky barrier height (0.98 eV) at 0 K and the temperature coefficient of the barrier ($-3.8 \times 10^{-4} \text{ V K}^{-1}$) agree with those reported for the IrSi phase implying that some reaction between Ir and Si has taken place. © 1997 American Vacuum Society. [S0734-211X(97)00804-4]

I. INTRODUCTION

Infrared detectors using iridium silicide Schottky barriers have attracted technological interest because they hold the promise of extending the detection range up to $12 \mu\text{m}$.¹ Iridium and its metallic silicides (IrSi, IrSi₃) are the metal-semiconductor (MS) contacts that present the highest Schottky barrier height (SBH) on *n*-type silicon, and thus the lowest SBH on *p*-type silicon.

Iridium reacts with silicon to form at least ten different phases,² although only three phases have been clearly identified, IrSi, IrSi_{1.75}, and IrSi₃³ in thin films. These three silicide compounds do not form sequentially as other silicides do. In fact several phases can exist at the same time. Studies of the electrical behavior of Ir_xSi_y/Si Schottky diodes have been carried out for contacts obtained by different procedures.^{1,4,5} These studies try to relate the measured SBH of the diodes with the existence at the MS interface of a particular phase, that depends on the heat treatment. Little has been published⁶ on the electrical behavior of the as-deposited iridium diodes, i.e., in the first stages of the silicide formation. Furthermore, these studies are limited to the determination of the SBH at room temperature, and therefore presents a very incomplete picture of the current transport across the barrier.

In this work we present a study of the Schottky diodes formed by the deposition of an iridium film on *n*-type silicon held at 150 °C with no further heat treatment. It will be shown that the diodes exhibit a highly ideal behavior at room temperature indicating that current transport is governed by thermionic emission of electrons over the top of the barrier. Deviations from this behavior at low temperature can be explained by the existence of a distribution of Schottky barrier heights at the iridium/silicide interface that results from the formation of a polycrystalline interlayer between the iridium

film and the silicon substrate in the early stages of silicide growth.^{7,8}

II. EXPERIMENT

Iridium silicide Schottky barrier diodes were prepared on (111) silicon wafers with a high-resistivity epilayer on a Sb doped substrate with a low resistivity ($0.05 \Omega \text{ cm}$). The $12 \mu\text{m}$ thick epilayer was doped with P to a resistivity of $4.2 - 5 \Omega \text{ cm}$. Following a standard RCA cleaning process, a thermal oxide 400 nm thick was grown at 1100°C on both sides of the wafers. Standard lithography procedures were used to define square windows in the front side of the wafer with side lengths of 100, 200, and $400 \mu\text{m}$. The oxide was etched during a 3 min immersion in a buffered hydrofluoric acid solution— NH_4F 4 vol, HF 1 vol—to expose the silicon in the regions where the diodes were to be formed. During this etching process, the oxide layer in the back side of the wafer was completely removed. The samples were then rinsed in deionized water, blown dry with nitrogen, and loaded immediately in a Joule effect evaporator where the back side was coated with a blanket layer of aluminum 300 nm thick. The back Ohmic contact was formed by annealing the samples at 450°C for 15 min in nitrogen.

After a 10 s immersion in a buffered solution of hydrofluoric acid— NH_4F 10 vol, HF 1 vol—to remove native oxides, the samples were rinsed in deionized water, blown dry with nitrogen, and loaded immediately in the electron gun evaporator where a pressure of $2 \times 10^{-7} \text{ Torr}$ was achieved by means of a conventional oil-diffusion pump with a baffle cooled by liquid nitrogen. The substrate holder consisted of a graphite susceptor in intimate contact with the sample. The temperature of the ensemble holder sample was sensed with a thermocouple embedded in the susceptor, that allowed to control the power applied to two halogen quartz lamps. Inside the chamber, the substrate holder was surrounded by a Meissner trap cooled with liquid nitrogen during the iridium deposition. Under these conditions, iridium films 20 nm

^{a)}Electronic mail: jleube@etsit.upm.es

thick were deposited at a rate of 0.05 nm s^{-1} on the wafers kept at 150°C under a vacuum better than 3×10^{-7} Torr. Several iridium slugs with a purity of 99.99 at. % were used as source material. The samples were allowed to cooldown under vacuum before being unloaded from the chamber. The iridium Schottky contacts were defined by a standard liftoff process. Finally, the diodes were cut and soldered into holders for electrical characterization.

The diodes were electrically characterized by the measurement of their I - V curves at different temperatures. To obtain the current-voltage characteristics the diodes were first cooled in a commercial cryostat. Then, a voltage was supplied by means of a Keithley-230 programmable voltage source while the current was measured with a Keithley-617 electrometer using a low impedance scheme. Both I - V measurements and temperature control were automated with the help of a personal computer via a general purpose interface bus (GPIB). The temperature was established by means of a commercial programmable controller with a calibrated Si diode mounted in the cold finger. A second diode measured the actual temperature of the sample. Following each temperature step, a delay was built into the program to allow the temperature difference between both sensors to become less than 0.3 K. Under such steady state conditions, a voltage was applied to the diode. A delay of a few seconds between subsequent steps was built into the program to ensure steady state conditions for the measurement of the current at low temperatures. Up to 50 readings were taken for each voltage step in order to increase the signal-to-noise ratio and extend the range of current measurements down to 2×10^{-13} A.

III. RESULTS AND DISCUSSION

A. Room temperature measurements

Figure 1 shows a semilogarithmic plot of the I - V characteristics measured for diodes of different areas at room temperature. To describe the forward I - V characteristics of the IrSi/Si Schottky diodes the well known thermionic emission equation is used,⁹ modified to take into account the series resistance effect:

$$I = I_{\text{sat}} \exp\left(\frac{qV_a}{nkT}\right) \left[1 - \exp\left(\frac{-qV}{kT}\right) \right],$$

$$I_{\text{sat}} = A_d A^{**} T^2 \exp\left(\frac{-q\Phi_B}{kT}\right), \quad (1)$$

$$V_a = V - IR_{\text{series}}.$$

Here A^{**} is the modified Richardson constant ($112 \text{ A/cm}^2 \text{ K}^2$ for n -type silicon¹⁰), A_d the diode area, n the ideality factor commonly used to measure the deviation of practical diodes from the ideal thermionic model, R_{series} the series resistance of the diode, q the electronic charge, k the Boltzmann's constant, T the temperature, and Φ_B the Schottky barrier height (SBH) that can be calculated from each I - V characteristic assuming that A^{**} and A_d are known.

In terms of the unidimensional model of the thermionic emission theory, discarding the low variations due to the

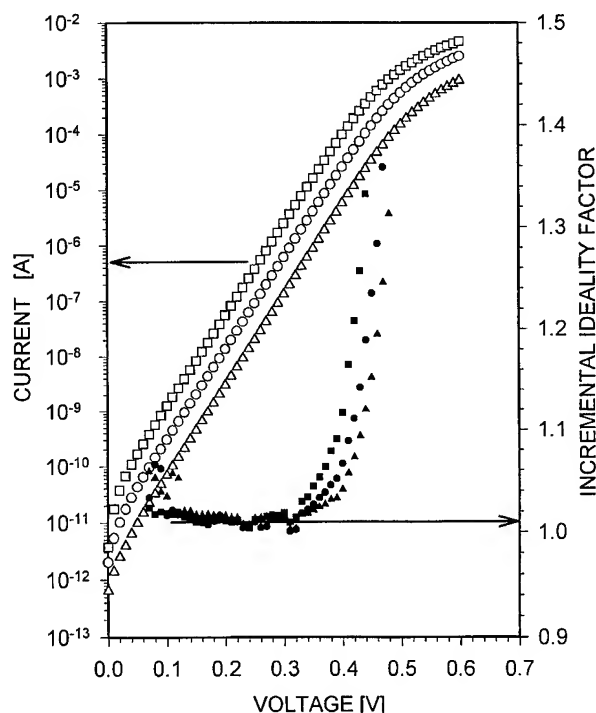


FIG. 1. Iridium/silicon as-deposited samples. Open symbols corresponds to the semilogarithmic plot of I - V characteristics at room temperature for three samples of $(100 \mu\text{m} \times 100 \mu\text{m})$, $(200 \mu\text{m} \times 200 \mu\text{m})$, and $(400 \mu\text{m} \times 400 \mu\text{m})$. Closed symbols correspond to the calculated incremental ideality factors for the diodes.

image-force lowering of the barrier,⁴ the ideality factor must remain close to unity independently of the applied bias and the measurement temperature,⁹ leading to a linear I - V semilogarithmic characteristic in the bias range where the series resistance effect can be neglected and for $V > 3kT/q$.

The straight-line behavior of the I - V characteristics represented in Fig. 1 over almost eight orders of magnitude is remarkable and suggests that the dominating current transport mechanism is thermionic emission. This fact was verified by the calculation of the quasi-Fermi level position relative to the bottom of the conduction band⁴ as a function of applied bias. The difference between the quasi-Fermi level position in the bulk and at the interface is less than 0.002 eV for all values of applied bias, confirming that the quasi-Fermi level remains flat throughout the depletion region and, therefore, that thermionic emission dominates the transport process. In the linear portion of the curves, the currents for a given voltage range scale well with the areal ratio of the diodes indicating that no peripheral effects take place.

Every parasitic transport mechanism (carrier recombination in the depletion layer of the semiconductor, tunnel effect, or field assisted tunnel effect)⁹ or the existence of Schottky barrier spatial distribution inhomogeneities¹¹⁻¹³ appear as current excesses in the semilogarithmic I - V characteristics of the diode. These excess currents can be detected easily by means of the representation of the incremental ideality factor variations as a function of the applied bias and/or

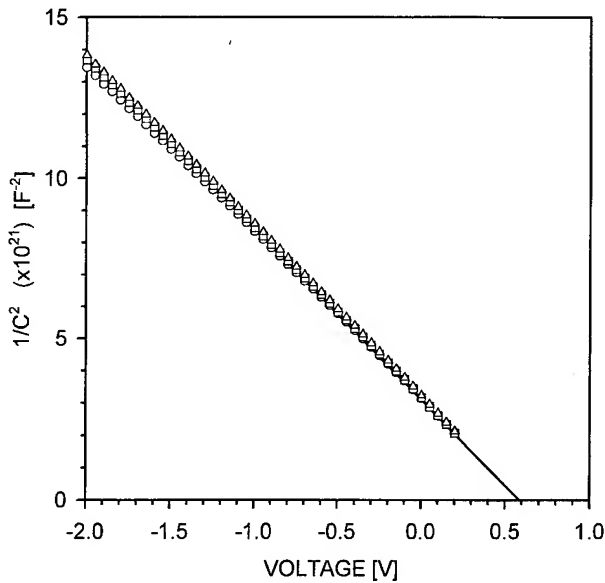


FIG. 2. Iridium/silicon as-deposited samples. $1/C^2-V$ characteristics at room temperature for three ($400\ \mu\text{m} \times 400\ \mu\text{m}$) samples at 298 K. $\Phi=0.607+0.263+0.0257=0.896\ \text{eV}$.

the measurement temperature. The incremental ideality factor n_i can be calculated by means of ⁹

$$n_i \approx \left[\frac{kT}{q} \left(\frac{\partial \ln(I)}{\partial V} \right) \right]^{-1} = \left[\frac{kT}{q} \left(\frac{\ln(I_{i+1})}{V_{i+1} - V_{i-1}} \right) \right]^{-1} \quad \left(V > \frac{3kT}{q} \right). \quad (2)$$

Thus, for an ideal Schottky barrier diode the value of n_i must: (a) be less than unity for $V < 3kT/q$ because of the effect of the pre-exponential term in the $I-V$ characteristics, (b) must be equal to unity in the middle bias range where the exponential term is dominant, and (c) increase with the applied bias where the series resistance effect dominates. Every parasitic current contribution appears as an increase of the ideality factor in the bias range where its effect is appreciable.

Figure 1 also shows the incremental ideality factor variation (n_i) obtained from the experimental data in each point as a function of the applied bias for the samples considered. No dependence of the incremental ideality factor with applied bias is observed in the voltage range where the effect of series resistance can be neglected. The calculated value for all samples agrees well with the theoretical value⁴ of 1.01 and the average effective barrier height calculated for these diodes is 0.88 eV.

$C^{-2}-V$ curves for the larger diodes (Fig. 2) were obtained at room temperature at a frequency of 1 MHz with an HP-LF-4192A impedance analyzer. The barrier height measured by this method is 0.90 eV. This value agrees fairly well with the values obtained from the room temperature forward $I-V$ characterization taking into account that the image-force lowering of the barrier¹⁴ is estimated to be 0.015 eV for

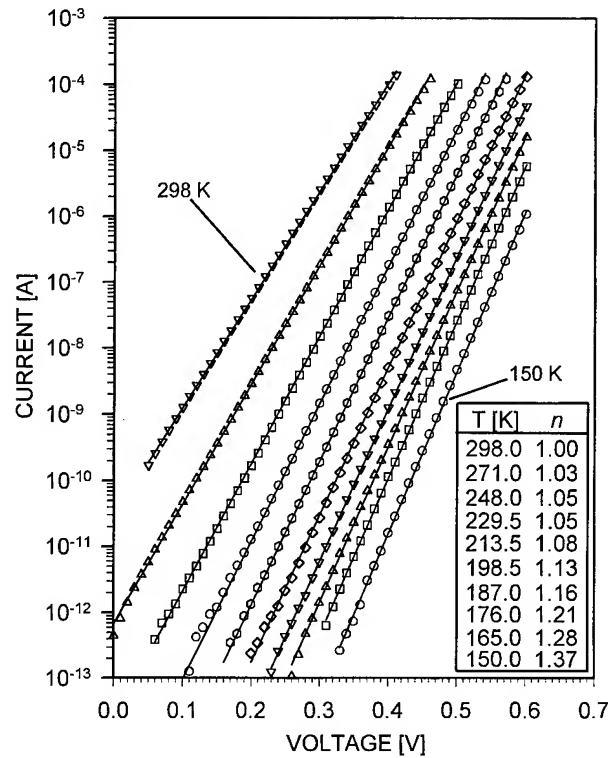


FIG. 3. Iridium/silicon as-deposited samples. $I-V-T$ characteristics of the diodes under study. The inset presents the calculated ideality factor for each measurement temperature.

our lightly doped epitaxial silicon. From the slope of this plot we can calculate the concentration of ionized donors in the substrate.

From the analysis presented above it is clearly seen that the Ir/Si Schottky diodes presented show nearly ideal room-temperature $I-V$ characteristics.

B. $I-V-T$ measurements

Figure 3 shows the $I-V$ curves measured in the temperature range from 150 to 300 K. The straight-line behavior in the semilogarithmic plot shows no dependence of ideality factor with applied bias for all temperatures. The ideality factors calculated for each temperature from the inverse of the slope of the regression line, as shown in the inset of Fig. 3, indicate that there is a strong variation of the ideality factor with temperature and thus a departure from simple thermionic emission theory.

Deviations of the ideality factor from unity can arise from several effects, like the presence of an interfacial oxide layer, the tunneling of electrons across the barrier, or recombination current in the depletion zone of the diodes. High resolution electron microscopy (HREM) micrographs of the diodes under study demonstrate that no interfacial oxide layer exists at the Ir/Si interface.^{7,8} However, these micrographs reveal the presence of a polycrystalline interlayer $\sim 3\ \text{nm}$ thick at the Ir/Si interface which is reasonably sharp. The presence of such an interfacial layer with no oxygen contamination was also suggested by SIMS analysis performed

on similar samples.¹⁵ Tunnel effects can be ruled out because of the lightly doped substrate and the temperature range investigated. Furthermore, there is no evidence of diffusion spikes in HREM micrographs, that would lead to appreciable tunneling currents. Finally, the recombination current in the depletion zone of the diode can also be discarded. The existence of such a current would lead to a saturation current larger than that of thermionic emission with an ideality factor close to two.⁹ The superposition of the recombination current and thermionic emission current would result in a non-linear plot of the I - V characteristics with an appreciable curvature in the low bias range, which have not been observed in our diodes. Therefore, none of the three mechanisms exposed can explain the strong temperature dependence of the ideality factor and its independence of the applied bias.

The value of the Schottky barrier height (Φ_{B0}) extrapolated to 0 K and the Richardson constant can be obtained from the slope of the plot of the saturation current divided by the square of the temperature versus the reciprocal temperature (Richardson plot). If the temperature variation of the barrier Φ_B can be modeled by^{16,17}

$$\Phi_B(T) = \Phi_{B0} + \alpha T, \quad (3)$$

the plot of I_{sat}/T^2 vs $1/T$ must be a straight line with a slope given by $-q\Phi_{B0}/kT$ and a Y axis intercept given by $(A_d A^{**} - \alpha/k)$. For the correct determination of Φ_{B0} the calculated value for the Richardson constant must coincide with its theoretical value. Therefore, the Richardson plot acts as a mirror of the interface homogeneity because, if there is any spatial fluctuation of the SBH at the interface¹⁸ the resulting plot is not linear and yields an underestimated or overestimated value of the Richardson constant.

Recent studies demonstrate that the electrical behavior of the iridium silicide Schottky barrier diodes under study can be explained by the existence of a spatially inhomogeneous Schottky barrier.¹¹⁻¹³ It has been proved that current-voltage characteristics presenting both, voltage independent and voltage dependent ideality factors, as well as temperature dependent n idealities, can be theoretically simulated by considering a nonuniform distribution of SBH at the MS interface and a unique conduction mechanism that is thermionic emission over the barrier.

The Richardson plot of I_{sat}/T^2 vs T^{-1} (Fig. 4) shows a very short linear region in the high temperature range, and strong deviations at low temperatures correlated to the departure from unity of the ideality factor. Since the experimental points do not lie in a straight line, it is not possible to determine the extrapolated SBH at zero bias and 0 K (Φ_{B0}) as the simple thermionic emission theory predicts. An approach to the value of Φ_{B0} could be obtained considering only the higher current values where the ideality factors approach unity, yielding 0.85 eV. The obtained Φ_{B0} value is lower than the SBH value obtained at room temperature in contrast with theory predictions [Eq. (3)], demonstrating that this method cannot be applied unless the Richardson plot is perfectly linear in a broad range of temperature.

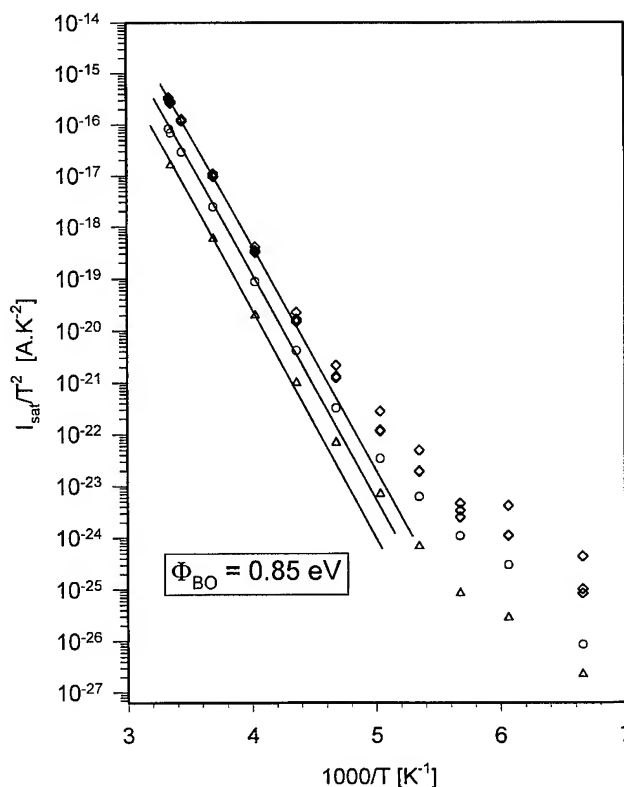


FIG. 4. Richardson plot (I_{sat}/T^2 vs $1000/T$) for five samples of different area. Three samples of ($400 \mu\text{m} \times 400 \mu\text{m}$), one of ($200 \mu\text{m} \times 200 \mu\text{m}$) and one of ($100 \mu\text{m} \times 100 \mu\text{m}$). (\diamond) $400 \mu\text{m} \times 400 \mu\text{m}$; (\circ) $200 \mu\text{m} \times 200 \mu\text{m}$; (\triangle) $100 \mu\text{m} \times 100 \mu\text{m}$.

It has been recently demonstrated that the SBH potential distribution at the MS interface depends on the microcrystalline properties of the interface rather than on the bulk properties of the constituents.¹⁹ For Schottky barrier diode characterization Werner *et al.*¹⁷ have proposed the use of the flatband barrier Φ_{bfp} [Eq. (4)] to compare experiments with theory, with Φ_{bfp} defined by

$$\Phi_{bfp} = n\Phi_B - (n-1) \frac{kT}{q} \ln\left(\frac{N_c}{N_d}\right), \quad (4)$$

where n is the ideality factor, Φ_B the Schottky barrier height at zero voltage calculated from I - V characteristics, N_c the density of states in the conduction band, and N_d is the ionized donor density. The SBH temperature coefficient and the value extrapolated to 0 K can be obtained from the plot of the flat-band barrier versus the temperature.

Figure 5 shows the variations of flatband SBH as a function of the measurement temperature assuming for simplicity that both the donor concentration (N_d) and the density of states in the conduction band (N_c) remain constant in this temperature range. The linear temperature dependence with temperature of the flat-band SBH can be clearly appreciated according to theoretical predictions. The temperature coefficient (α) for the SBH is $-3.8 \times 10^{-4} \text{ V K}^{-1}$ and the value

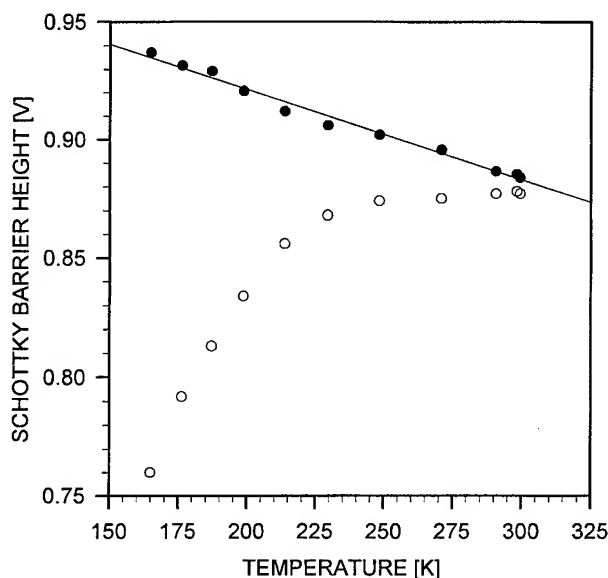


FIG. 5. Flat-band Schottky barrier height (Φ_{BFP}) variation as a function of the measurement temperature. Temperature coefficients are directly obtained from the linear regression. (●) flat-band SBH; (○) apparent SBH.

extrapolated at 0 K is 0.98 eV. Both (Φ_{B0} and α) agree very well with previously reported values for the monosilicide phase IrSi.¹⁷

The electrical behavior of the *as-deposited* diodes can be related with the existence of the already mentioned polycrystalline interlayer that appears in the Ir/Si interface reported previously⁷ according to HREM observations. This interlayer appears like a very faulted area with a composition that cannot be exactly determined and formed of grains of different size and orientation corresponding to the first stages of silicide formation.

IV. CONCLUSIONS

Ir/Si metallizations with no further heat treatment are excellent high barrier Schottky diodes at room temperature. Room temperature SBH determined by I - V and C - V characterization agree if one takes into account the image force reduction of the barrier.

The study of the behavior as a function of temperature indicates that there is a high degree of inhomogeneity in the SBHs at the MS interface. Flat-band SBH variation analysis implies that there is some IrSi formation at the MS interface as a consequence of a partial reaction between iridium and silicon.

The properties of the SBH for Schottky diodes are strongly influenced by the nature of the few monolayers immediately in contact to the silicon substrate.

ACKNOWLEDGMENT

This work was supported by the CICYT through Grant No. MAT93-0040.

- ¹B.-Y. Tsaur, C. K. Chen, and B. A. Nechay, IEEE Electron Device Lett. **11**, 415 (1990).
- ²M.-A. Nicolet, in *VLSI Electronics: Microstructure Science* (Academic, San Diego, 1983), Vol. 6, Chap. 6.
- ³S. Petersson, J. Baglin, W. Hammer, F. d'Heurle, T. S. Kuan, I. Ohdomari, J. de Sousa Pires, and P. Tove, J. Appl. Phys. **50**, 3357 (1979).
- ⁴M. Wittmer, Phys. Rev. B **42**, 5249 (1990).
- ⁵B.-Y. Tsaur, M. M. Weeks, R. Trubiano, P. W. Pellegrini, and T.-R. Yew, IEEE Electron Device Lett. **9**, 650 (1988).
- ⁶I. Ohdomari, K. N. Tu, F. M. d'Heurle, T. S. Kuan, and S. Petersson, Appl. Phys. Lett. **33**, 1028 (1978).
- ⁷C. Ballesteros, T. Rodríguez, F. J. Jiménez-Leube, and M. Clement, J. Appl. Phys. **77**, 5173 (1995).
- ⁸T. Rodríguez, H. Wolters, M. Fernández, A. Almendra, M. F. da Silva, M. Clement, J. C. Soares, and C. Ballesteros, Appl. Surf. Sci. **73**, 182 (1993).
- ⁹E. H. Rhoderick, in *Metal Semiconductor Contacts* (Clarendon, Oxford, 1980).
- ¹⁰S. M. Szé, *Physics of Semiconductor Devices*, 2nd. ed. (Wiley, New York, 1981), Chap. 5.
- ¹¹J. H. Werner and H. H. Güttler, J. Appl. Phys. **69**, 1522 (1991).
- ¹²J. P. Sullivan, R. T. Tung, M. R. Pinto, and W. R. Graham, J. Appl. Phys. **70**, 7403 (1991).
- ¹³R. T. Tung, Phys. Rev. B **45**, 13509 (1992).
- ¹⁴E. H. Rhoderick and R. H. Williams, *Metal-Semiconductor Contacts* (Oxford Science, London, 1988).
- ¹⁵J. M. Blanco, J. J. Serrano, F. J. Jiménez-Leube, T. Rodríguez, and H. Wolters, Vacuum **45**, 1121 (1994).
- ¹⁶J. Tersoff, Phys. Rev. B **32**, 6968 (1985).
- ¹⁷J. H. Werner and H. H. Güttler, J. Appl. Phys. **73**, 1315 (1993).
- ¹⁸Y. P. Song, R. L. van Meirhaeghe, W. H. Laflère, and F. Cardon, Solid-State Electron. **29**, 633 (1986).
- ¹⁹R. T. Tung, J. Vac. Sci. Technol. B **11**, 1546 (1993).

Linewidth dependence of stress relaxation and microstructural change in passivated Al(Cu) lines*

I.-S. Yeo and P. S. Ho^{a)}

Center for Materials Science and Engineering, University of Texas at Austin, Austin, Texas 78758-1100

S. G. H. Anderson and H. Kawasaki

Advanced Products Research and Development Laboratory, Motorola, Austin, Texas 78721

(Received 12 July 1996; accepted 30 May 1997)

Relaxation of thermally induced stresses in passivated line structures can lead to void formation, and its rate is strongly correlated with microstructural change in the lines. A bending beam technique was used to measure the thermal stress relaxation behavior of passivated Al (1 wt % Cu) line structures with 3, 1, and 0.5 μm linewidths as a function of temperature up to 15 h. Our data indicate that stress relaxation in Al(Cu) lines depends on line geometry and temperature, and exhibits log (time) kinetics, which is consistent with a thermally activated dislocation glide mechanism. The observed behavior can be explained by a combined result of the driving force due to the effective shear stress and kinetics due to thermally activated mass transport. Scanning electron microscope and transmission electron microscope techniques were used to study stress-induced voiding as a function of aging time and temperature. A strong correlation between stress relaxation rate and void density was observed for 1 μm wide lines, but not for 0.5 μm wide lines. A lower void density of 0.5 μm wide lines suggests that the bamboo grain structure may influence the mass transport and retard void formation. © 1997 American Vacuum Society.

[S0734-211X(97)15404-X]

I. INTRODUCTION

Since the observation of stress-induced voiding phenomenon in Al interconnects in 1984,^{1,2} many studies of this problem have been performed. It is generally accepted that relaxation of tensile stresses in Al interconnects constitutes the driving force for void formation.³ Stress voiding endurance of layered interconnects and its relation to stress relaxation have been reported;^{4,5} however, a thorough investigation of stress relaxation behavior for Al interconnects is desirable for understanding the mechanism(s) of void formation and for improving Al interconnect reliability. It is clear that stress evolution and relaxation depend on both anneal temperature and line geometry. To understand this problem, we investigated the linewidth dependence of stress relaxation and the relationship between stress relaxation and stress-induced voiding. This article summarizes the results of this study, including some results briefly reported previously.^{6,7}

The stress relaxation kinetics of passivated Al (1 wt % Cu) line structures with 3, 1, and 0.5 μm linewidths were investigated using a bending beam technique. Test structures were annealed at 400 °C for 30 min, and cooled to temperatures suitable for measurement. It should be noted that stress relaxation after cooling from an elevated temperature is of the greatest interest because stress-induced voiding results from the relaxation of tensile stresses. Stress relaxation was measured at 250, 200, and 150 °C for 15 h.

We limited the stress relaxation measurements between 250 and 150 °C during cooling from 400 °C because stress relaxation is a combined result of the driving force and mass

transport kinetics. When the stress relaxation temperature is low, then the driving force that comes from the thermal stress is high, but mass transport is slow, hence stress relaxation is very slow. When the stress relaxation temperature is high, mass transport is fast, but the driving force is low, so stress relaxation is again very slow. Consequently, there exists a temperature range where the combination of the stress driving force and the mass transport rate is optimized for a stress relaxation measurement. This temperature range was found to extend from 150 to 250 °C for this study. Within this temperature range, the observed stress relaxation kinetics are linewidth and temperature dependent.

The statistics of void formation were measured as a function of aging time and temperature using a scanning electron microscope (SEM). The microstructure of passivated lines and void morphology were investigated using a plan view transmission electron microscope (TEM). The relationship between stress relaxation and stress-induced voiding will be discussed.

II. EXPERIMENT

Tetraethylorthosilicate (TEOS) oxide and boron phosphorus doped silicate glass (BPSG) films of 100 and 600 nm thickness, respectively, were deposited on a 625 μm thick Si substrate to prevent reaction between Al and the Si substrate. This was followed by a deposition of 600 nm Al (1 wt % Cu) film and a 25 nm TiN anti-reflective coating. The Al film was patterned using standard photolithographic and etch procedures to form periodic arrays of Al lines with nominal widths of 3, 1, and 0.5 μm . These line structures were defined on different portions of the same wafer in order to

*No proof corrections received from author prior to publication.

^{a)}Electronic mail: paulho@mail.utexas.edu

obtain identical sample preparation conditions for different linewidth structures. The line samples were then annealed for 30 min at 400 °C to stabilize the microstructure. A bilayer passivation of 400 nm phosphorus doped silicate glass (PSG) and 700 nm plasma-enhanced silicon nitride (PEN) was then deposited on the line samples and on the oxidized Si substrate. The stress measurements for the line samples were supplemented by similar measurements for blanket films of bilayer passivation and metal with and without the bilayer passivation. The stress versus temperature characteristics of the blanket films are required for deducing thermal stresses of passivated lines.⁸

The samples were diced to 3 mm wide by 45 mm long strips with periodic lines aligned parallel or perpendicular to the long direction of the strips. Each sample included only one type of line (orientation and width). These strips were thinned to a thickness of $\sim 400 \mu\text{m}$ to improve the measurement sensitivity. Also, the backside of the sample was polished to a mirror finish in order to improve the reflectance for the laser beam. The polished backside of the samples was used as a laser reflecting surface in order to avoid any artificial effect due to diffraction of light from the periodic line structures on the frontside of the wafer.

These strips were then inserted into the stress measurement system, as described in Ref. 9. After a 30 min anneal at 400 °C in the system, stress relaxation was measured at temperatures of 250, 200, and 150 °C for 15 h. After measurement, the passivation films and metal lines were chemically removed in order to determine the film-free wafer curvature. These measured radii of curvature were then converted to the composite stresses using an extension of Stoney's equation for anisotropic films.¹⁰ It is worth noting that, in line samples, stresses along and across the lines are interdependent as measured by the respective radii of curvature through Poisson's effect; thus both parallel and perpendicular lines to the sample beam should be measured. The principal stresses in passivated metal lines were then deduced from these composite stresses.⁸

Since the microstructure of the Al(Cu) plays an important role in controlling the stress and its relaxation behavior, TEM was used to analyze its grain structure. For this purpose, the passivated line samples for microstructure study were annealed at 400 °C for 30 min to stabilize the microstructure of the Al(Cu) and to simulate the proper condition for stress relaxation measurement. The passivation layer was removed by reactive ion etching before thinning the sample to $\sim 100 \mu\text{m}$. Dimpling to $\sim 10 \mu\text{m}$ and ion milling at 5 kV was performed until an electron-transparent region was reached. A JEOL 1200 EX scanning TEM (STEM) was used for the microscopy work. The grain size distribution of the passivated lines was determined from plan view TEM micrographs.

For the stress-induced voiding study, a TEM and a SEM were used to examine samples aged for 2 years at room temperature and then annealed at 250, 200, and 150 °C for 24 h. These samples were similar to those used for stress relaxation studies as they were prepared in the same wafer

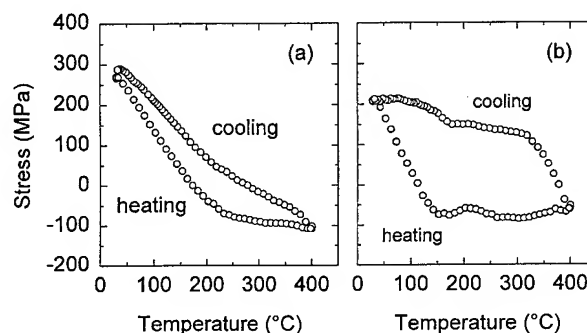


FIG. 1. Stress vs temperature curves for (a) 0.6 μm thick unpassivated Al (1 wt % Cu) film and (b) 0.6 μm thick Al (1 wt % Cu) film passivated with a PEN/PSG bilayer film.

lot. In order to obtain stress voiding statistics for 1 and 0.5 μm wide lines, 10 positions were selected at random and examined by a field emission SEM at 10k \times magnification. For 3 μm wide lines, 5 positions were selected at random and examined at 5k \times magnification. For the TEM void morphology study, samples were prepared using the dimpling and ion milling processes described earlier.

III. RESULTS

A. Blanket films

Figures 1(a) and 1(b) show typical stress-temperature curves of unpassivated and passivated blanket Al (1 wt % Cu) films. These two films exhibit quite different stress versus temperature behavior, with the unpassivated film showing less stress hysteresis compared to the passivated film. This is caused primarily by their different stress behaviors during cooling. During cooling, the unpassivated film exhibits a smooth continuous increase in stress with decreasing temperature as is dictated by the thermal mismatch between the film and the substrate. There is some yielding during cooling in the unpassivated film, as is evidenced by a smaller stress versus temperature slope than that during heating from 25 to 200 °C. In comparison, the passivated film shows a sharp increase in stress during cooling from 400 to 320 °C. This is characterized by a stress versus temperature slope equal to that during initial heating from 25 to 150 °C, thus indicating an elastic behavior upon initial cooling from 400 °C. This is followed by a gentle increase during subsequent cooling to 25 °C, revealing a distinct yield behavior from 320 to 25 °C. This behavior has been investigated¹¹ and can be attributed to an effect caused by an increase in the average Al(Cu) grain size during the deposition of the passivation layer.

Figures 2(a) and 2(b) show the stress relaxation data of the unpassivated and the passivated blanket films at different annealing temperatures. The stress relaxation data, when plotted as a function of log (time), show a linear relationship after annealing for more than 20 min. This is consistent with previous results,¹² and suggests stress relaxation via a thermally activated dislocation glide mechanism.

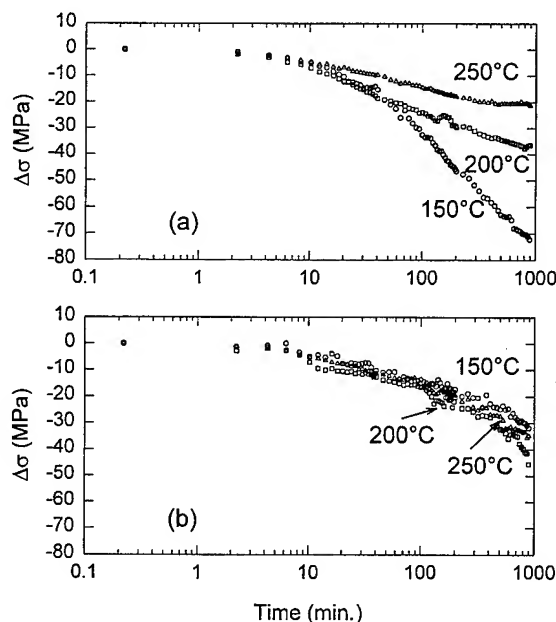


FIG. 2. Stress relaxation behavior of (a) unpassivated film and (b) passivated film. Here $\Delta\sigma = \sigma(t) - \sigma(t=0)$.

Stress relaxation by a thermally activated dislocation glide mechanism can be described by

$$\sigma(t) = \sigma(t=0) - \sigma^* \ln\left(1 + \frac{t}{\tau}\right), \quad (1)$$

where $\sigma(t=0)$ is the initial stress at the stress relaxation temperature T , σ^* is an effective stress relaxation rate that is determined from the slope of the σ vs $\log(\text{time})$ plot, and τ is a relaxation time.^{13,14} We note that the term σ^* in Eq. (1) represents the stress change per decade of time and has a unit of stress; thus it is a convenient parameter to compare the stress relaxation behavior of different films and lines as a function of temperature.

Stress relaxation data for the blanket films were fit by Eq. (1) in order to obtain the effective stress relaxation rate σ^* as shown in Fig. 3. It is clear that the unpassivated and passivated films show quite different stress relaxation behavior. For the unpassivated film, the stress relaxation rate increases monotonically as relaxation temperature decreases. For the passivated films, the stress relaxation rate shows a peak at 200 °C but overall it is a weak function of temperature. (Although the temperature dependence of the relaxation rate for the passivated film is weak, the measured variation is statistically significant.) Note that the temperature dependence of the stress relaxation rate shown in Fig. 3 for the passivated and the unpassivated films resembles the respective stress versus temperature curves during cooling in Fig. 1. The reason will be discussed later.

B. Passivated lines

The stress versus temperature curves for a 1 μm wide passivated Al(Cu) line are shown in Fig. 4. Stresses along, across, and normal to the lines (σ_x , σ_y , and σ_z) were de-

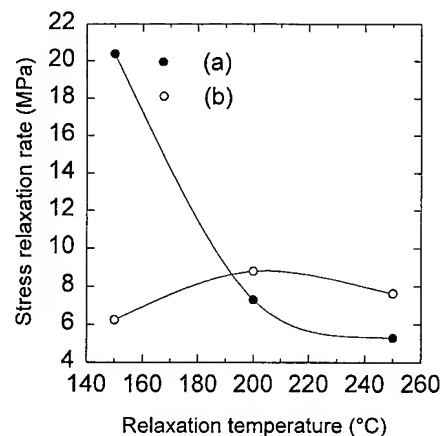


FIG. 3. Stress relaxation rate as a function of temperature for (a) 0.6 μm thick unpassivated and (b) passivated Al (1 wt % Cu) film (PEN/PSG bilayer passivation).

termined using a micromechanical analysis described in Ref. 8. These curves demonstrate that large compressive and tensile stresses are generated in passivated Al lines as they are heated and cooled. The thermal stresses are larger than those observed for blanket films of comparable thickness, and are triaxial because of the presence of the passivation layer. Stress changes as a function of temperature are the greatest along the length of the line, σ_x , and least along the normal direction, σ_z .

Stress relaxation data for 1 μm wide lines at 250 °C for 15 h are shown in Fig. 5. Examination of the three curves indicates several distinct features. First, the rate of stress relaxation is highest in the x (along the line) direction, then followed by y (across the line) and z (normal to the line) directions. Second, the stress relaxation data show a linear $\log(\text{time})$ relationship after an initial period of 20–30 min. This suggests that the time-dependent stress changes in pas-

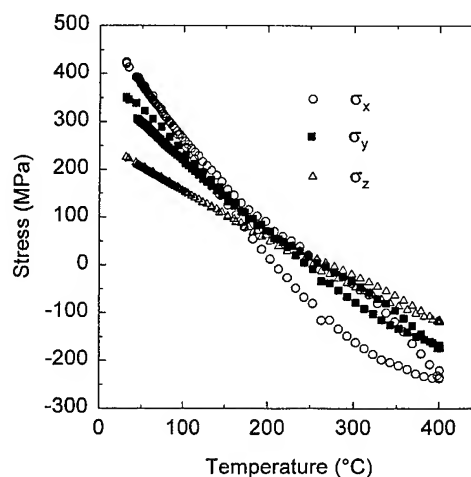


FIG. 4. Stress vs temperature curve of 1 μm wide Al (1 wt % Cu) lines passivated with a PEN/PSG bilayer film.

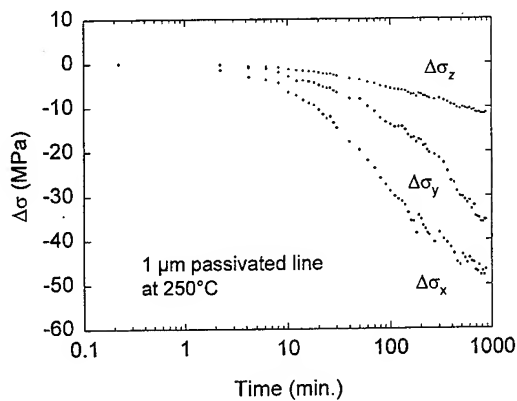


FIG. 5. Stress relaxation behavior of 1 μm wide passivated line at 250 $^{\circ}\text{C}$ during cooling. Here $\Delta\sigma = \sigma(t) - \sigma(t=0)$.

sivated lines may also be controlled by a thermally activated dislocation glide mechanism during the first 15 h of isothermal annealing.

Stress relaxation data for the principal directions at temperatures of 150, 200, and 250 $^{\circ}\text{C}$ were fit to Eq. (1) for the 3.0, 1.0, and 0.5 μm wide lines in order to obtain the effective stress relaxation rate σ^* . Figure 6 shows that the stress relaxation rate depends both on the linewidth and the anneal temperature. Similar to the results shown in Fig. 5, stress relaxation along the line (σ_x^*) is fastest, followed by the stresses across the line (σ_y^*), with stress relaxation normal to the line (σ_z^*) being slowest. The overall stress relaxation can be evaluated by considering the sum of the individual relaxation rates. On that basis, the 1 μm wide line results shown in the middle panel exhibit a maximum at 200 $^{\circ}\text{C}$, whereas the stress relaxation behaviors of 3 μm and 0.5 μm wide lines are different. As shown in the leftmost panel, the overall stress relaxation for 3 μm wide lines is largest at 150 $^{\circ}\text{C}$, then decreases continuously from 200 to 250 $^{\circ}\text{C}$. For 0.5 μm wide lines, the trend is opposite as is illustrated in the rightmost panel where the overall stress relaxation increases with increasing temperature. The overall stress relaxation

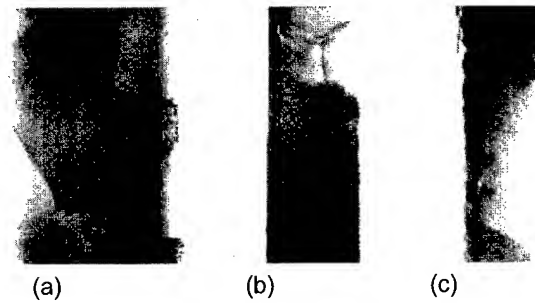


FIG. 7. Microstructure of passivated Al (1 wt % Cu) line for (a) 3, (b) 1, and (c) 0.5 μm wide lines, respectively.

characteristics can be understood by considering the interplay of the driving force which originates from thermal stresses and the kinetics of mass transport and depend on temperature and grain structure. As the relative effect of these two factors changes for different linewidths, the temperature dependence of the stress relaxation rate varies accordingly. This point will be discussed later in connection with data on thermal stress and grain structure as a function of linewidth in Sec. IV.

C. Microstructure and stress-induced voiding of passivated lines

Stress relaxation is known to depend on the microstructure of the lines. Grain structures of the passivated line samples after annealing at 400 $^{\circ}\text{C}$ for 30 min were examined using a TEM and are shown in Fig. 7. The 3 μm wide lines show a uniform polycrystalline structure, whereas 1 and 0.5 μm wide lines exhibit near bamboo and true bamboo structures, respectively. The grain size distributions of the passivated 3, 1, and 0.5 μm wide lines shown in Fig. 8, follow a log-normal distribution. The median grain sizes of 3, 1, and 0.5 μm wide lines are 1.69, 0.88, and 0.83 μm , respectively. The median grain sizes decrease as the linewidth decreases, indicating that grain growth is inhibited by line edges. This

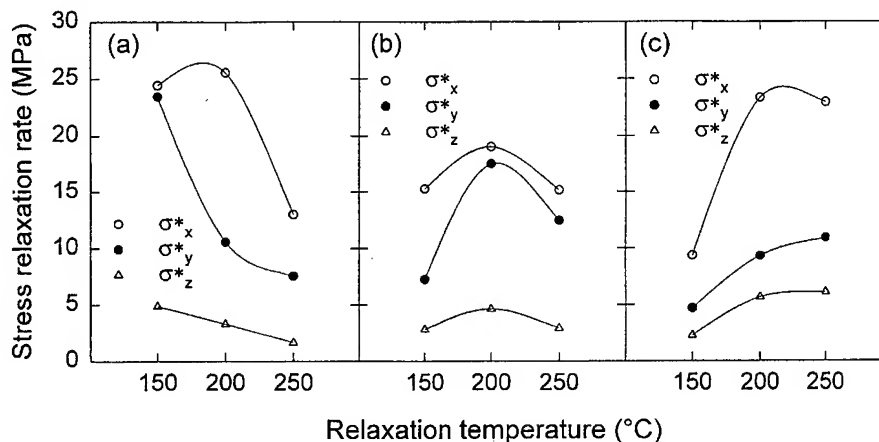


FIG. 6. Stress relaxation rate of passivated Al (1 wt % Cu) line as a function of temperature for (a) 3, (b) 1, and (c) 0.5 μm wide lines, respectively.

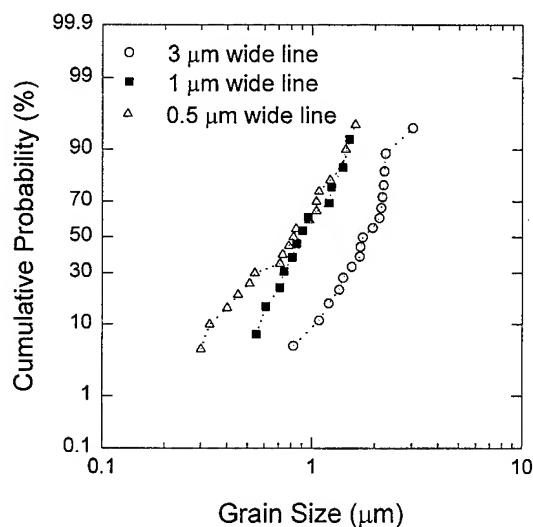


FIG. 8. Grain size distribution of the passivated Al (1 wt % Cu) line for different linewidths.

observation is consistent with a previous study by Kwok *et al.* who observed also a decrease in grain size with decreasing linewidth.¹⁴

Figure 9 shows void morphology in the passivated line samples stored for 2 years at room temperature and then annealed at 200 °C for 24 h. Voids are located at intersections between grain boundaries and line edges, and are predominantly wedge shaped. The typical size of the voids is in the range of 0.1–0.2 μm. SEM micrographs in Fig. 10 show typical void distributions for samples of 3.0, 1.0, and 0.5 μm wide lines. The void densities which are determined from the SEM micrographs and presented by open-circle data points, are correlated with the corresponding stress relaxation rates and the results are presented in Fig. 11. [In Fig. 10(c), the 0.5 μm wide lines are separated by lines with a regularly varying linewidth that are introduced for different purposes. The voids appearing in those lines are not counted in the data presented in Fig. 11(b).] The void density for 1 μm wide lines shows a good correlation with the relaxation rate although the data for 0.5 μm wide lines do not. Figures 10 and

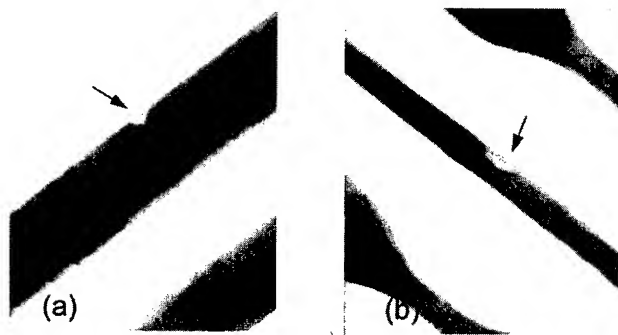


FIG. 9. Void morphology for (a) 1 and (b) 0.5 μm wide lines. Voids are located at the intersection of the grain boundary and line edge, and are predominantly wedge shaped.

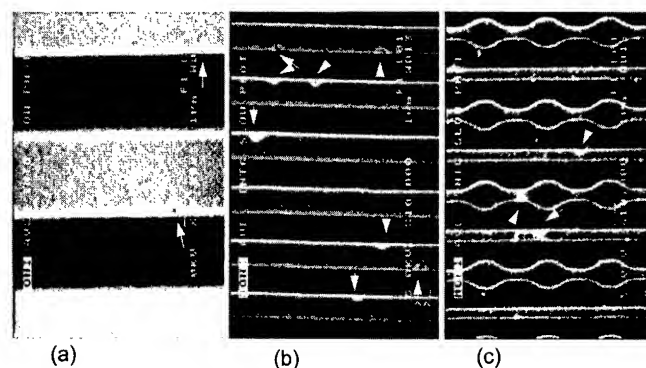


FIG. 10. Void distribution for (a) 3, (b) 1, and (c) 0.5 μm wide lines after a 24 h anneal at 200 °C. Arrows indicate the void location.

11 reveal that 0.5 μm wide lines generally have a lower void density than the 1 μm wide lines.

IV. DISCUSSION

A. Linewidth dependence of stress relaxation

In general, the stress relaxation rate depends both on driving force (thermal stress) and mass transport kinetics. The kinetics of mass transport are slow at low temperatures and fast at high temperatures. In contrast, the driving force which comes from thermal stresses during cooling is usually low at high temperatures and high at low temperatures, as illustrated in Figs. 1 and 4. In general, thermal stresses depend on temperature, linewidth, line height, and passivation. When plastic yield is involved, thermal stress will depend on the grain structure. Mass transport is a function of temperature and grain structure. For lines with near-bamboo (1 μm) or bamboo (0.5 μm) grain structures, mass transport will occur through the coupling of grain boundary diffusion with interfacial diffusion. In such cases, mass transport will also depend on line geometry. Overall, stress relaxation behavior depends on thermal stress, temperature, and line geometry.

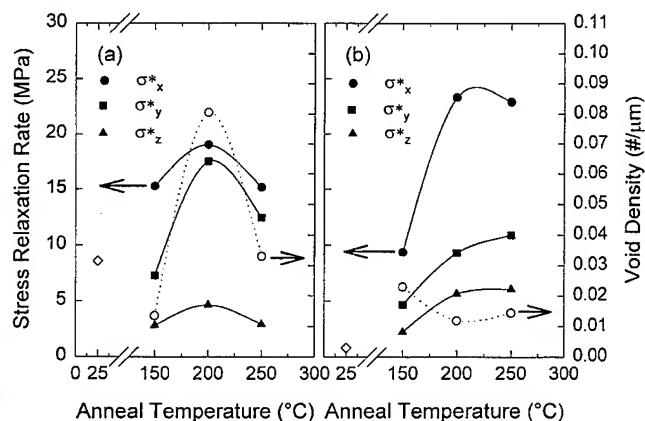


FIG. 11. Comparison of stress relaxation and void densities for (a) 1 and (b) 0.5 μm wide lines. The open circles represent data for void density which refers to the scale on the right side as indicated by the arrows.

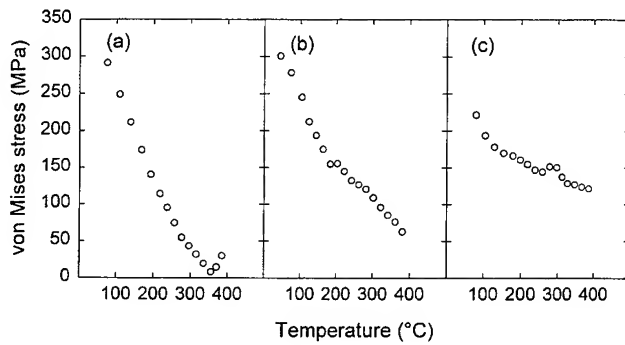


FIG. 12. The von Mises stress of passivated Al (1 wt % Cu) lines as a function of temperature for (a) 6, (b) 1, and (c) 0.5 μm wide lines, respectively.

Early studies showed that stress-induced voiding, and hence stress relaxation, exhibited a maximum near 180 °C for 2–3 μm wide lines.^{15,16}

When stress relaxation kinetics are consistent with thermally activated dislocation glide, the driving force for stress relaxation is a shear component of thermal stress instead of thermal stress itself.¹⁷ To investigate the nature of the driving force for stress relaxation, the von Mises yield stresses were examined for each of the blanket and line samples as a function of temperature.⁶ The effective shear stress, which is a driving force for dislocation glide, is defined by von Mises yield criterion as

$$\sigma_Y = \frac{1}{\sqrt{2}} [(\sigma_x - \sigma_y)^2 + (\sigma_y - \sigma_z)^2 + (\sigma_z - \sigma_x)^2]^{1/2}, \quad (2)$$

where σ_x , σ_y , and σ_z are the principal stress components.¹⁸ It should be noted that, for the case of thin films, $\sigma_x = \sigma_y$ and $\sigma_z = 0$, so the von Mises stress reduces to the measured stress, that is, $\sigma_Y = \sigma_x = \sigma_y$. Thus, for thin films, the measured stress during cooling is the driving force for stress relaxation. This is not applicable to passivated line structures because of the triaxial stress state.

From Fig. 1(a), the driving force (the measured thermal stress during cooling) of the unpassivated film for stress relaxation shows a large monotonic increase with decreasing temperature. Thus, if the temperature dependence of mass transport is not as large as the change in the thermal stress, the stress relaxation behavior would be dominated by the driving force. This seems to be the case for the unpassivated film which exhibits a temperature dependence of stress relaxation resembling that of the thermal stress during cooling. In contrast, the passivated film in Fig. 1(b) shows a small stress increase during cooling between 250 and 150 °C, so the effect of mass transport for stress relaxation probably plays a more important role. As a result of the combination of the driving force and mass transport kinetics, the stress relaxation rate of the passivated film shows a peak at 200 °C, although the overall temperature dependence of the relaxation rate is quite weak.

The von Mises stresses of passivated lines were calculated from the principal stresses of 6, 1, and 0.5 μm wide passi-

vated Al lines and are shown in Fig. 12. The stresses were determined by an x-ray diffraction technique¹⁹ for the same type of samples prepared by the Advanced Products Research and Development Laboratory (APRDL) of Motorola. Stresses measured by the x-ray technique during cooling were used to calculate the von Mises stresses because of the inaccuracy in determining the zero stress position for bending beam methods.^{8,20}

The von Mises stresses as a function of temperature show a consistent trend of increase with decreasing temperature with the fastest rate of increase for the widest line of 6 μm . Unfortunately, there was no stress-temperature x-ray diffraction data for the 3 μm wide lines to compare directly with the stress relaxation data. However, since the von Mises stresses exhibit an overall consistent trend as a function of temperature for linewidths ranging from 0.5 to 6 μm , we assume that the von Mises stress of 3 μm wide lines will also exhibit a rapid change as the temperature decreases, similar to the 6 μm lines. For narrow lines, the change in von Mises stresses as a function of temperature is considerably slower. Indeed, between 250 and 150 °C during cooling, the von Mises stresses of the 0.5 μm wide lines are almost constant.

These results can be used to understand the linewidth dependence of the stress relaxation results shown in Fig. 6. For wider lines (3 μm wide), the driving force (the von Mises stress at the temperature) for stress relaxation increases significantly as temperature decreases, and it controls the overall kinetics of stress relaxation. Thus, the stress relaxation rate of 3 μm wide lines is largest at 150 °C. In contrast, for 0.5 μm wide lines, the change in von Mises stresses between 250 and 150 °C during cooling is small [Fig. 12(c)], so the kinetics of stress relaxation is controlled by mass transport. As a result, the overall stress relaxation is fastest at the highest temperature (250 °C). For 1 μm wide lines, the temperature dependence of the relaxation rate comes about from the combination of driving force and mass transport. As a result, the rate is maximized at an intermediate temperature of 200 °C.

It is interesting to note that the average grain sizes for the 3, 1, and 0.5 μm lines are 1.69, 0.88, and 0.83 μm , respectively, and the 3 μm lines have a uniform polycrystalline structure, while the 1 and 0.5 μm lines have near-bamboo and bamboo grain structures. Since grain boundary diffusion dominates mass transport in the temperature range of the stress relaxation experiments, the rate of mass transport will depend on the size and distribution of the grains. For the 3 μm line, its uniform polycrystalline structure will provide a grain boundary network to promote mass transport for stress relaxation. This results in an overall higher stress relaxation rate than the narrow lines with near-bamboo and bamboo grain structures. For the narrow lines, interfacial diffusion will contribute to mass transport by coupling to the grain boundary diffusion. At this time, it is not clear to what extent interfacial diffusion contributes to mass transport, although bamboo grains are known to reduce the overall mass transport under electromigration. However, the stress relaxation

rate is quite high for the 0.5 μm lines at 250 °C, indicating that the role of interfacial diffusion for stress relaxation may be different from electromigration. This problem remains to be investigated.

B. Stress relaxation and voiding

Stress relaxation is generally thought to consist of two stages involving plastic yielding and diffusional flow. In the first stage, the principal stresses are relaxed to a nearly hydrostatic state. This is manifested in our studies where the stress relaxation rate is the highest for the largest principal stress σ_x along the line direction, and decreases in the order of across the line y and normal to the line z directions. The relative magnitude of the stress relaxation rate will result in a stress state approaching hydrostatic. In the second stage, diffusion flow occurs under nearly hydrostatic stress conditions, which leads to void formation. In actual cases, plastic yielding and diffusional flow do not take place in distinct sequence; instead, both would occur simultaneously and contribute to stress relaxation although with different kinetics. Initially, stress relaxation via plastic flow is greater than that via diffusion, and then diffusional flow contributes more at longer times.

All of the stress relaxation results, such as those shown in Figs. 2 and 5, can be fit to a model for thermally activated dislocation glide. Note that dislocation glide is a volume conserving process, so hydrostatic stresses cannot be relaxed. However, the data in Fig. 5 show that all three principal stress components decrease during isothermal annealing. This indicates that a non-volume conserving mechanism is also involved in stress relaxation, suggesting that a diffusion flow process occurs concurrently with plastic flow to relax thermal stresses. For such a mechanism, void formation, a volume change due to Al_2Cu precipitation, and passivation deformation during aging could be possible candidates.^{21,22}

Precipitation and growth of Al_2Cu during aging have been observed by some investigators.^{23–25} Since the atomic volume of Al_2Cu is larger than the sum of the volumes of two Al and one Cu atoms in an Al matrix, precipitate growth can account for some of the stress relaxation measured in the line. This would be consistent with the reduction observed in all three stress components, as shown in Fig. 5; however, the amount of relaxation is too small to explain the results observed.²¹ Besser has considered the effect of passivation deformation during isothermal annealing.²¹ Since the passivation layer is not perfectly rigid, it can deform in response to the stresses in Al lines to reduce all three stress components, and thus explain the observed reduction of hydrostatic stresses. However, such a stress relaxation process is athermal, contradicting the observed results which show a strong dependence on anneal temperature.

Void formation seems to be a likely candidate to explain the non-volume conserving stress relaxation. Indeed, a clear correlation between the stress relaxation rate and the void density (or void formation rate) is observed for 1 μm wide lines, although such a correlation cannot be established for 0.5 μm wide lines due to the low void density. Further SEM

observation of voids for 0.5 μm wide lines is underway to search for such a correlation. The lower void density observed in 0.5 μm wide lines (Figs. 10 and 11) can be attributed to a change in grain structure. As shown in Fig. 7, 1 μm wide lines have a near-bamboo structure with some embedded polycrystalline clusters, whereas 0.5 μm wide lines show an almost perfect bamboo structure. These observations support the idea that bamboo grains retard void formation, which is similar to the effect of reduced electromigration damage in bamboo grain line structures.

For interconnect reliability assessment, it is important to establish a correlation between stress relaxation and void formation. Although voids in passivated lines are believed to form primarily by relaxation of hydrostatic tensile stresses,³ a correlation between void formation and stress relaxation was observed in this experiment over a 15 h period. This result implies that void formation should depend on both linewidth and anneal temperature in a manner similar to that observed for stress relaxation. Consequently, the optimum condition for stress voiding should depend on both line geometries and anneal temperature, suggesting that predicting the linewidth dependence of stress voiding based on results obtained at a fixed temperature might lead to severe errors. The present study is limited to the first stage of 15 h of stress relaxation, which seems too short to have significant void growth.²² In order to study the stress-induced voiding in detail and clear manner, long-term stress relaxation experiments would be necessary.

In order to study stress-induced voiding, a resistometric method has been widely used.²² It is worth noting that the stress relaxation measurement can be complementary to the resistometric method. The resistometric method is more sensitive to large voids extending across the entire line,²⁶ whereas the stress relaxation method may be more sensitive to the total void volume. Further experimental work is underway to confirm the correlation of the stress relaxation behavior and stress-induced voiding in the long term.

V. CONCLUSIONS

Stress relaxation in passivated Al (1% Cu) lines was measured by a bending beam technique. We observed that stress relaxation follows a log (time) kinetics, consistent with a thermally activated dislocation glide mechanism. The relaxation rate was found to depend on line geometry and stress relaxation temperature. The observed behavior can be explained by a combined result of driving force due to the effective shear stress and kinetics due to thermally activated mass transport. The void density for different linewidths was examined by SEM as a function of temperature. A strong correlation between the stress relaxation rate and void density was observed for 1 μm wide lines, but not for 0.5 μm wide lines. A lower void density of 0.5 μm wide lines suggests that the bamboo grain structure may influence the mass transport and retard void formation.

ACKNOWLEDGMENTS

This work was supported by Semiconductor Research Corporation (Contract No. 94-MJ-353), Texas Advanced Technology Program (Grant No. 156), and Motorola, APRDL, Austin. The authors would like to thank to Dr. A. P. Clark and Professor S. Saimoto of Queen's University, Canada, for the x-ray stress measurements data. The authors also would like to thank Dr. P. B. Ghate for valuable discussions.

- ¹J. Klema, R. Pyle, and E. Domagoue, IEEE Proc. Int. Relat. Phys. Symp. **22**, 1 (1984).
- ²J. Curry, G. Fitzgibbon, Y. Guan, R. Muollo, G. Nelson, and A. Thomas, IEEE Proc. Int. Relat. Phys. Symp. **22**, 6 (1984).
- ³H. Okabayshi, Meas. Sci. Eng. **11**, 191 (1993), and references contained therein.
- ⁴N. Matsunaga and H. Shibata, IEEE Proc. Int. Relat. Phys. Symp. **32**, 256 (1994).
- ⁵Y. Kawano and T. Ohta, IEEE Proc. Int. Relat. Phys. Symp. **33**, 353 (1995).
- ⁶I.-S. Yeo, S. G. H. Anderson, C.-N. Liao, H. Kawasaki, and P. S. Ho, Mater. Res. Soc. Symp. Proc. **338**, 281 (1994).
- ⁷I.-S. Yeo, C.-N. Liao, H. Kawasaki, and P. S. Ho, presented at the 3rd Workshop on Stress-Induced Phenomena in Metallization, Stanford, CA, 1995.
- ⁸I.-S. Yeo, S. G. H. Anderson, P. S. Ho, and C.-K. Hu, J. Appl. Phys. **78**, 953 (1995).
- ⁹I.-S. Yeo, M.S. thesis, University of Texas at Austin, 1993.
- ¹⁰J.-H. Jou and L. Hsu, J. Appl. Phys. **69**, 1384 (1991).
- ¹¹P. A. Flinn, D. S. Gardner, and W. D. Nix, IEEE Trans. Electron Devices **34**, 689 (1987).
- ¹²M. A. Korhonen, C. A. Paszkiet, R. D. Black, and Che-Yu Li, Scr. Metall. Mater. **24**, 2297 (1990).
- ¹³H. Schroeder, U. Burges, and W. Schilling, in *2nd Workshop on Stress-Induced Phenomena in Metallization, Austin, TX, 1993*, edited by P. S. Ho, C. Y. Li, and P. A. Totta, AIP Conf. Proc. No. 305 (AIP, New York, 1993).
- ¹⁴T. Kwok, C. Y. Ting, and J.-U. Han, IEEE V-MIC Conference, 1985, pp. 83-89.
- ¹⁵J. W. McPherson and C. F. Dunn, J. Vac. Sci. Technol. B **5**, 1321 (1987).
- ¹⁶J. T. Yue, W. P. Funsten, and R. V. Taylor, IEEE Proc. Int. Relat. Phys. Symp. **23**, 126 (1985).
- ¹⁷H. J. Frost in *Materials Reliability in Microelectronics II*, edited by C. V. Thompson and J. R. Lloyd [Mater. Res. Soc. Symp. Proc. **265**, 3 (1992)].
- ¹⁸G. E. Dieter, *Mechanical Metallurgy*, edited by S. I. Metric (McGraw-Hill, London, 1988), p. 76.
- ¹⁹A. P. Clark, S. Saimoto, and P. S. Ho, in Ref. 13.
- ²⁰S. G. H. Anderson, I.-S. Yeo, P. S. Ho, A. P. Clark, and S. Saimoto, presented at the 1994 MRS Spring Meeting, Materials Reliability in Microelectronics, 1994.
- ²¹P. R. Besser, Ph.D. dissertation, Stanford University, 1993.
- ²²M. G. Fernandes, H. Kawasaki, J. L. Klein, D. Jawarani, R. Subrahmanyan, T.-K. Yu, and F. Pintchovski, in Ref. 13, pp. 153-164.
- ²³C. Kim and J. W. Morris, Jr., J. Appl. Phys. **72**, 1837 (1992).
- ²⁴E. G. Colgan and K. P. Rodbell, J. Appl. Phys. **75**, 3423 (1994).
- ²⁵T. Nemoto and T. Nogami, IEEE Proc. Int. Relat. Phys. Symp. **32**, 207 (1994).
- ²⁶L. A. Miller and A. K. Stamper, 1995 VMIC Conference, 1995, p. 369.

Probeless voltage contrast using a focused ion beam for opens and shorts defect isolation of ultralarge scale integration technologies

Kenneth J. Giewont,^{a)} Douglas B. Hunt, and Klaus M. Hummler^{b)}

IBM Microelectronics Division, Advanced Semiconductor Technology Center, Route 52, Hopewell Junction, New York 12533

(Received 1 March 1996; accepted 18 April 1997)

Defect isolation in large yield mazes for accurate and timely physical analysis continues to be a critical part of process development for 0.4 and 0.25 μm technologies. Shrinking dimensions have challenged common approaches such as light emission microscopy (LEM), optical beam induced current (OBIC), liquid crystal, and mechanical probing to the point where alternate techniques are needed. This article describes simple but powerful defect isolation techniques using probeless voltage contrast in a focused ion beam (FIB) tool. Fail modes to which the techniques are applicable include deep trench capacitor polysilicon to substrate leakage, gate conductor to self-aligned contact shorts, metal line opens, via chain opens, and metal line to line shorts. When possible, OBIC or LEM are initially used to determine the approximate fail location followed by precise isolation using the FIB. Typical examples of each fail mode will be shown. The basic principle employs the positive focused ion beam to charge floating structures which tend to attenuate the secondary electron image. This provides a suitable potential difference between grounded and ungrounded structures which can be easily observed in the electron imaging mode. Using this phenomenon along with leveraging the ability of the FIB to cut and deposit replaces the need for a probe. In addition, once a defect is isolated it can be precisely marked and then unlayered or cross sectioned *in situ*. This technique has provided the only practical solution to isolate defects in some situations while reducing the average turn around time for feedback to process engineering. © 1997 American Vacuum Society. [S0734-211X(97)06104-0]

I. INTRODUCTION

Timely and accurate failure analysis of large yield mazes has been an important component of yield learning for the IBM/SIEMENS/TOSHIBA 64 Mb¹ and 256 Mb² DRAM programs at IBM Microelectronics. The large critical area combined with simplistic structures enables quick identification of yield detractors without full product build and test. Common yield mazes include metal line comb and serpentine structures, via chains, deep trench capacitor arrays, and gate conductor mazes with self-aligned contacts. These structures can be built as a short loop routing or as part of a fully integrated process. Identification of failure mechanisms in these structures has become increasingly difficult as linewidths continue to shrink and critical areas need to be increased. Conventional defect isolation techniques such as light emission microscopy (LEM), optical beam induced current (OBIC),³ and liquid crystal are still useful for leakage mechanisms but often lack spatial resolution to pinpoint the fail site to a single cell. OBIC has the best spatial resolution but, if the defect is located in the middle of a large array, tedious procedures must be performed to ascertain the location. Isolation of opens is commonly done using mechanical probing or e-beam techniques such as voltage contrast or specimen current. The problems associated with mechanical probing are obvious when linewidths approach 0.25 μm . Electron beam techniques are usually sufficient in finding

most opens, however to accurately record the fail location often requires tedious counting back to a reference point or recording stage movement. When a cross section of the defect is required and the defect is not optically visible (as is often the case for sub-micron defects) the e-beam techniques become impractical. In contrast, the focused ion beam (FIB) allows immediate sectioning or highly accurate marking of the fail site for later physical analysis. Finding shorts in large metal line mazes has always been the most challenging problem since few techniques can accurately pinpoint the fail site. If liquid crystal does not work, the defect can be very roughly located by measuring the voltage divider created by the short between the comb and serpentine. Unfortunately,

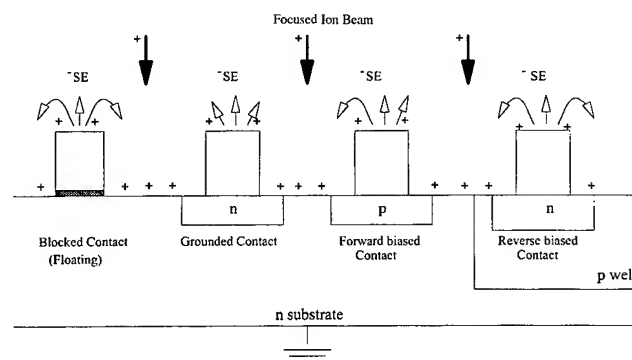


Fig. 1. Diagram showing how secondary electrons are attenuated as the surface is charged by the ion beam, except in the case of the grounded contact and forward biased contact.

^{a)}Electronic mail: KGIEWONT@VNET.IBM.COM

^{b)}Also with Siemens Components Inc.

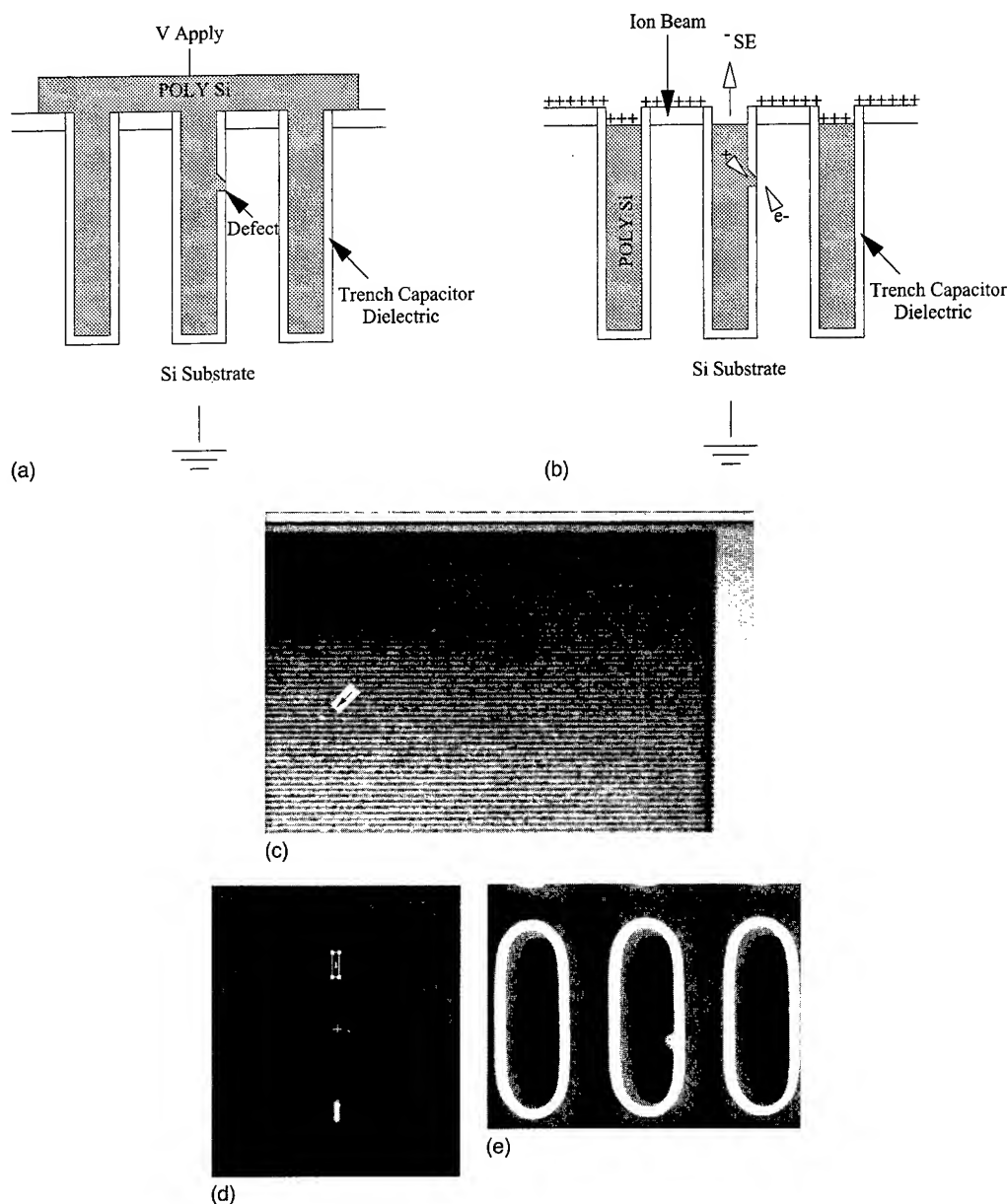


FIG. 2. (a) Trench capacitor test structure. (b) Trench shorted to substrate allows charge to dissipate, allowing more secondary electrons to be detected. (c) OBIC image showing approximate location of trench capacitor short in a 64 000 trench test structure. (d) FIB image in electron mode showing brighter contrast of trench capacitor which is shorted. (Box indicates where reference mark will be milled.) (e) Trench poly removed showing capacitor dielectric breakdown.

this only works if a single defect is present and at best only isolates the defect to a region consisting of several metal lines. The FIB tool has been found to be a very effective defect isolation tool by imaging the voltage contrast induced by the interaction of the ion beam and the electrical behavior of the open or short. Most defect mechanisms in the mazes mentioned above can be quickly and accurately located and then permanently marked for subsequent analysis. Leveraging the ability of the FIB to cut and deposit metal lines allows isolation of metal line opens and shorts without the need for *in situ* probes. Similar techniques have been applied using electron beams.⁴ However, these techniques are limited because they usually require active probes and, most impor-

tantly, do not allow *in situ* circuit modification. For deep trench capacitor leakage and gate conductor leakage in large mazes, LEM or OBIC assist in approximating the location of the fail. The FIB is then used to pinpoint the fail site to a single cell. Significant improvements in turnaround time and no defect found (NDF) rate have been realized using this technique. In addition, a method has been demonstrated for finding metal line shorts, which otherwise had been nearly impossible to find using conventional techniques. This article will discuss the theory of operation and will describe the isolation technique used for four practical applications. These include deep trench capacitor leakage, gate conductor leakage, metal line opens, and metal line shorts.

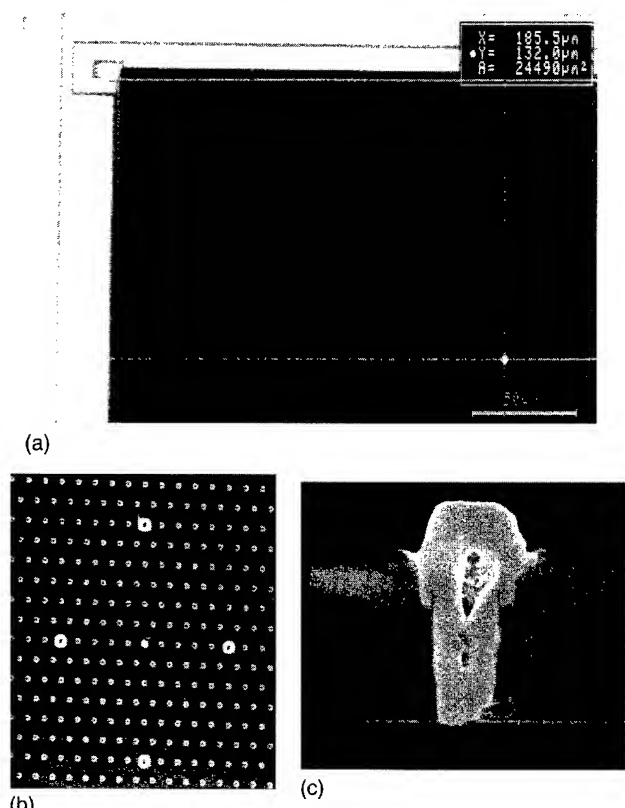


FIG. 3. (a) OBIC image showing approximate location of metal contact to poly Si gate short in a 500 000 contact test structure. (b) FIB image showing exact metal contact shorted to poly gate. Location has been "marked" for mechanical cross-section with milled boxes. (c) SEM micrograph of mechanically cross-sectioned contact to poly gate short.

II. THEORY OF OPERATION

When the gallium ion beam of a FIB tool strikes the surface of a conductor it produces a yield of greater than one secondary electron per ion. There is thus a net depletion of negative charge from the target unless it is grounded. Therefore, a floating conductor will tend to charge positive and recollect its own secondary electrons. If the ion beam strikes a grounded conductor the potential and thus the secondary emission will tend to stay relatively constant. As a consequence the floating conductor will appear darker in the secondary electron image than a similar grounded structure. Using this simple concept along with cutting and patching metal lines allows the FIB to become a powerful defect isolation tool. In the process of isolating a defect it is often necessary to ground or float a conductor. To ground a metal line a hole is milled to the silicon substrate near or on the line and is then filled with tungsten. Alternately, to float a conductor, the line is simply cut from the grounding path. A voltage contrast can also be imaged between conductors tied to *n*-type diffusions versus *p*-type diffusions. The positive ion beam will tend to reverse bias *n*-type diffusions and thus allow the conductor potential to rise above 1 V. In comparison, conductors tied to *p*-type diffusions will charge to forward bias the diffusion and thus limit the voltage to about 0.6 V. A third type of voltage contrast can be imaged by

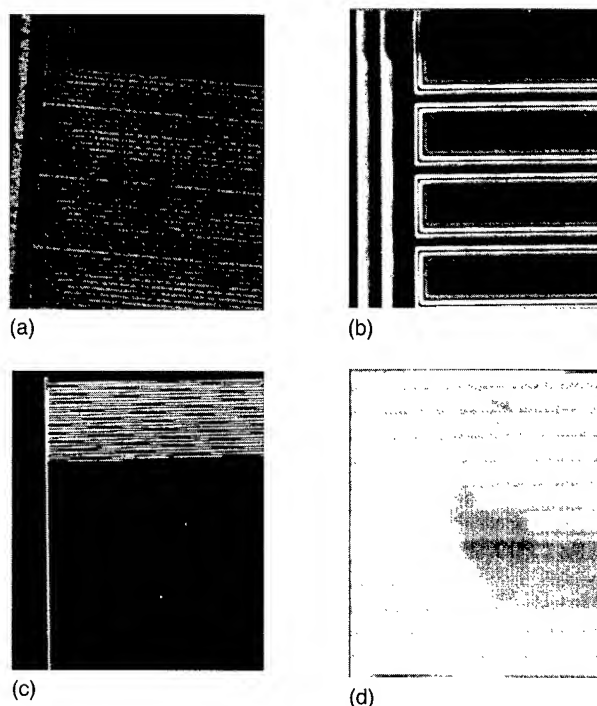


FIG. 4. (a) Metal line serpentine/comb test structure with 5 serpentines and 2 combs. Entire structure is built over Si oxide and is electrically floating. (b) Two of the serpentines have been grounded by milling a hole from the test pad down to the Si substrate. (c) Low magnification FIB image illustrating the brighter image obtained from the grounded serpentines on one side of a metal open. (d) Exact location of metal open is easily determined from contrast difference.

secondary electrons when a small conductor is shorted to a floating conductor with very large capacitance. An example is when an isolated contact to *n*-type diffusion is shorted to a very long gate conductor. Good contacts surrounding the shorted contact will charge positive as the diffusion is reverse biased thus decreasing their secondary electron emission. The shorted contact will only begin to charge positive once the large conductor is charged. This allows the shorted contact to be "effectively grounded" long enough to image its voltage contrast to the surrounding good contacts. Figure 1 illustrates the various types of voltage contrast described above. The examples which follow were done on a Micron 9000 FIB with typical beam conditions set at 30 keV with a 2 μ A extraction current. Aperture sizes varied between 100 and 400 μ m (corresponding to beam currents of 0.06–1.8 nA) depending on the application. It is desirable to operate with the 100 μ m aperture for higher resolution and to minimize beam damage. However, there are occasions when a larger aperture is needed to increase the sample current, which results in a larger potential difference between good and bad structures. One example is when trying to isolate high resistance contacts where higher beam current provides a larger voltage drop across the bad contact and thus yields a stronger voltage contrast.

III. APPLICATIONS

Examples are given where this technique was used to isolate defects as part of a 64 Mb DRAM development program.

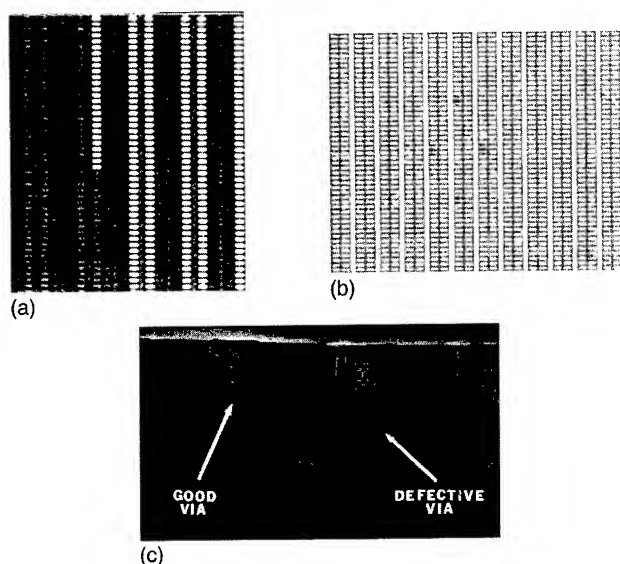


FIG. 5. (a) FIB image of a metal-via-metal interlevel chain open. One end of the chain has been grounded. (b) Optical micrograph showing fail "marked" for mechanical cross section. (c) SEM micrograph of mechanically cross-sectioned open via.

A. Deep trench capacitor leakage isolation

In order to accelerate yield learning in the 64 Mb program, wafers were periodically pulled from product lots and a special mask was used to connect large blocks of deep trench capacitor cells in parallel using the polysilicon fill as the conductor. These blocks of trenches could then be measured for polysilicon to substrate leakage across the trench insulator [Fig. 2(a)]. In order to isolate failing trenches for failure analysis, OBIC or copper decoration techniques are typically used [Fig. 2(c)]. Often these techniques allow one to approximate the location of the fail, but precise isolation to an individual failing trench is difficult. Even if one is successful in isolating a trench, it is difficult to record the location of that trench in a large structure. Using the FIB as a probeless voltage contrast tool, an individual trench can be isolated after getting an approximate location from one of the above techniques. The polysilicon is first recessed to electrically isolate the trenches using traditional wet etch or reactive ion etch (RIE) techniques. The chip is then placed in the FIB and imaged in the electron mode, causing the fail to appear brighter than the surrounding trenches [Figs. 2(b) and 2(d)]. The defect can then be marked by milling boxes nearby, and subsequently unlayered or cross-sectioned for scanning electron microscope (SEM) or transmission electron microscope (TEM) to determine the cause of fail [Fig. 2(e)].

B. Polysilicon gate to metal contact short isolation

The 64 Mb product cell uses a borderless contact to diffusion defined by the polysilicon gate. A productlike test site connects all the gates as a large serpentine and connects the metal contacts as another serpentine using the first level of metal. The two serpentes can then be tested for gate to

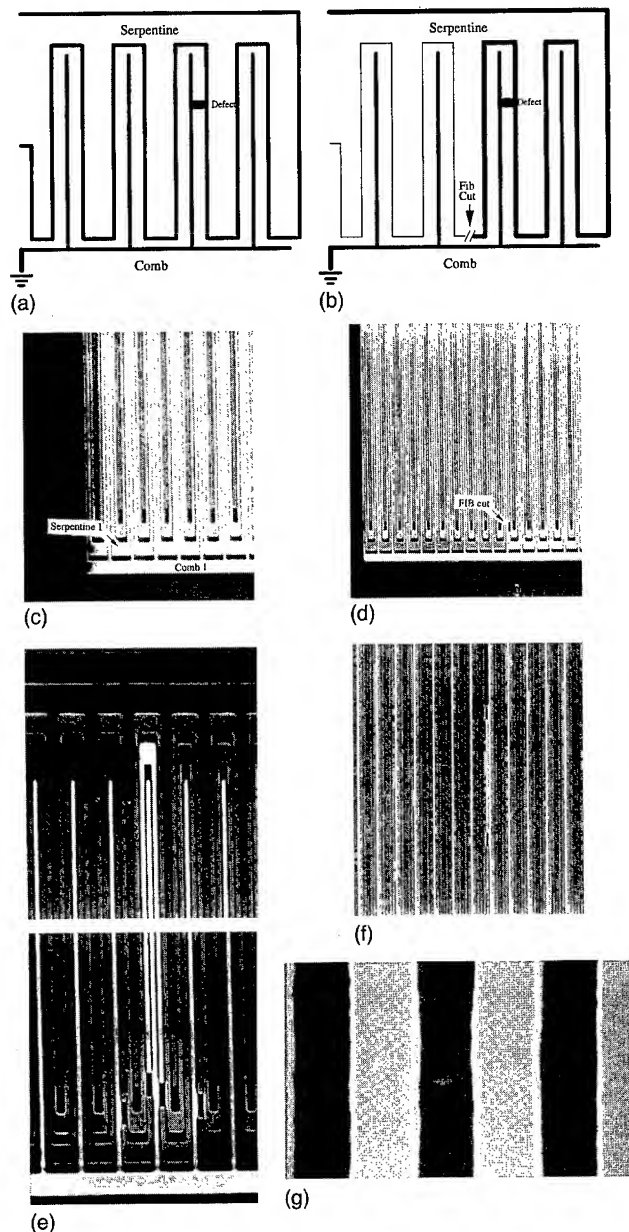


FIG. 6. (a) Typical comb serpentine metal shorts test site with comb grounded. (b) FIB cut causing left side of serpentine to be floating, thus reducing the secondary electron signal. (c) FIB image of a metal line serpentine/comb test structure. Comb 1 has been grounded, serpentine 1 is shorted to it. (d) Serpentine has been cut in the FIB. The left side is no longer shorted to the comb so it now appears darker. (e) Short has been isolated to one serpentine line. (f) Short has been isolated to a small section of the serpentine line. (g) SEM micrograph of metal short showing a thin Ti/TiN liner bridge.

metal contact shorts. To isolate the defect, OBIC was traditionally used because the contacts are connected to isolated diffusions. Isolating the defect to one contact can be tedious, especially if the fail site is located several hundred microns from any unique reference point. After OBIC is used to approximate the location of the fail [Fig. 3(a)], imaging the identified region in the FIB allows further isolation to a single contact. Since the failing contact is effectively grounded to the large poly gate serpentine, it will appear

much brighter than the surrounding contacts [Fig. 3(b)]. The contact can then be marked with milled boxes near the defect and subsequently unlayered or cross-sectioned to determine the cause of fail [Fig. 3(c)].

C. Metal conductor line and via chain opens

A number of yield structures with metal line serpentes and intermetal via chains are included in the development kerf of the 64 Mb product chip. Full chip test sites are also used to evaluate process changes in the back-end of line processing. The metal serpentes on these test sites can be quite large, some covering the equivalent to a 16 Mb area of the chip. Finding small metal opens becomes impractical using optical inspection. Via chain opens are usually not visible at all unless they are located first and unlayered or cross-sectioned. Using the FIB, metal opens can be isolated by grounding one end of the metal serpentine. This can be done by milling a hole in the metal probing pad deep enough to touch the Si substrate. The hole can then be filled with tungsten if the FIB is equipped with a tungsten deposition option. Often, by intentionally making a "rough" cut, i.e., using a large aperture with no further touch up, the tungsten deposition is not needed since enough metal has been redeposited in the hole to make the connection. The metal serpentine is now grounded and will appear bright in an electron mode image [Figs. 4(a) and 4(b)]. The part of the serpentine beyond the open is still floating and will appear darker. Using this difference in contrast, the defect can easily be isolated [Fig. 4(c) and 4(d)]. Similarly, an interlevel metal via chain open can be isolated in the same way. One end of the chain is grounded which causes the entire chain to appear brighter until the open is reached [Fig. 5(a)]. The failing via can then be marked and cross-sectioned or unlayered [Figs. 5(b) and 5(c)].

D. Metal conductor line to line shorts

Using the FIB probeless voltage contrast technique to find metal shorts has become extremely valuable for the 64 Mb program. Because of the large sizes of the serpentine mazes (see above), even isolating a short down to a region of a maze using the short as a voltage divider is not accurate enough. Many defects such as TiN stringers at the bottom of an Al line, or an assortment of underlying defects in the case of the metal damascene process⁵ cannot be seen visually. Therefore, defect isolation down to a very small area is necessary for subsequent SEM review. In a typical application, the comb of the maze is grounded by milling a hole in the

metal probing pad deep enough to touch the Si substrate. This causes the metal comb to appear bright in the electron imaging mode. If the adjacent metal serpentine is shorted to the comb, it will also appear bright since it is now grounded [Figs. 6(a) and 6(c)]. The serpentine is cut using the FIB to mill away a section of line. The portion of the serpentine shorted to the comb still appears bright, but the portion now floating is dark since it is no longer grounded [Figs. 6(b) and 6(d)]. By making successive cuts in the serpentine, the fail can be isolated to one row of the maze [Fig. 6(e)]. Cutting the serpentine in half each time, this method will require at most N cuts (where 2 to the power of N equals the number of serpentine rows) to locate one defect. If there is more than one defect present, they can all be isolated by continuing with the same procedure. By further cutting, the defects can be isolated to a small area [Fig. 6(f)]. As a bonus, it has been noticed that many embedded shorts are easier to image in the FIB than the SEM. Defects can now be marked for SEM imaging, cross-sectioning, or unlayering [Fig. 6(g)].

IV. CONCLUSIONS

Leveraging the FIB as a probeless voltage contrast tool has been shown to provide practical defect isolation techniques for opens and shorts in large yield structures. Since FIB tools are commonly installed at most semiconductor fabs, this technique gives the failure analysis engineer another defect isolation tool without additional capital investments. Enhancements to this technique may include probes in the FIB tool that would allow further flexibility if biasing is needed. In addition, the techniques described here have been successfully applied to DRAM product chips. The success rate for finding defects whose exact location cannot be deducted from a bit-fail map (bit line fails, word line fails, block fails) has been greatly improved.

ACKNOWLEDGMENT

The authors would like to acknowledge Mike Jennison's early involvement with this technique before he left IBM.

¹D. Kenney *et al.*, 1992 Symposium on VLSI Technology, Digest of Technical Papers, June 1992, pp. 14 and 15.

²G. Bronner *et al.*, *Proceedings of 1995 Symposium on VLSI Technology* (IEEE, New York, 1995), pp. 15 and 16.

³Keith C. Stevens and Thomas J. Wilson, *Microelectron. Eng.* **12**, 397 (1990).

⁴M. Macari *et al.*, *Proceedings of the Reliability Physics Symposium*, 1982, pp. 163-165.

⁵S. Roehl *et al.*, *Proceedings of the 9th International VLSI Multilevel Interconnection Conference*, 1992, pp. 22-28.

Surface roughness of nitrided (0001) Al_2O_3 and AlN epilayers grown on (0001) Al_2O_3 by reactive molecular beam epitaxy

Wook Kim,^{a)} M. Yeadon, A. E. Botchkarev, S. N. Mohammad, J. M. Gibson, and H. Morkoç^{b)}

Materials Research Laboratory and Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

(Received 13 August 1996; accepted 16 May 1997)

Nitridation of *c*-plane sapphire followed by deposition of a low temperature AlN or GaN buffer is commonly employed in the growth of GaN-based structures which have received a great deal of attention recently. In order to gain some needed insight, we undertook an investigation of nitridation of sapphire followed by the subsequent growth of AlN buffer layers in a reactive molecular beam epitaxy environment. Atomic force microscopy was used to characterize the surface roughness of samples after exposure to various nitridation conditions. Nitridation at higher temperatures was found to yield smoother substrate surfaces possibly due to smoothing of scratches introduced during substrate preparation. Incorporation of nitrogen into sapphire surfaces during the nitridation process was verified using x-ray photoelectron spectroscopy by observing the development of the N 1s peak with nitridation time. The surface roughness of AlN layers deposited on these nitrided surfaces was found to increase dramatically with thickness due to a significant coarsening of the surface topography. Surface roughness was found to decrease with increasing growth rate, the smoothest films being obtained with a growth rate of 140–200 nm/h at a substrate temperature of 800 °C.

© 1997 American Vacuum Society. [S0734-211X(97)13904-X]

I. INTRODUCTION

Group III nitrides have emerged as highly promising wide band gap semiconductors for both electronic and optoelectronic device applications.¹ Recent interest in these semiconductors has led to the development of high-efficiency blue light emitting diodes² and blue laser diodes.³ Significant progress has also been made in the development of electronic devices including GaN/AlGaIn modulation-doped field effect transistors⁴ (MODFETs). Prepared by reactive molecular beam epitaxy (RMBE), recent MODFETs yielded high transconductance (≥ 200 mS/mm) and drain-source current (≥ 600 mA/mm). Although SiC substrates exhibit a number of advantages over sapphire such as smaller lattice mismatch and better thermal conductivity, many researchers still obtain better results on sapphire, necessitating investigations such as the present one.

It has been observed that thin low temperature AlN (Ref. 5) or GaN (Ref. 6) buffer layers enhance the GaN crystalline quality grown on them in the metal organic chemical vapor deposition (MOCVD) process. Koide *et al.*⁷ suggested that, in MOCVD, the low-temperature-grown AlN buffer layer plays a significant role by providing sufficient nuclei for the alignment of the crystal axis to the *c*-axis, and by reducing the fluctuation of the crystal mosaic structure. When the AlN film is grown on a (1 $\bar{1}$ 02) oriented sapphire substrate after low-pressure nitridation at 1200 °C, the sapphire surface is chemically converted into a composite of Al, N, and O by the nitridation process, leading to a smoother surface mor-

phology of the film.⁸ A transmission electron microscopy study of the interface between AlN and (1 $\bar{1}$ 02) sapphire substrate also identified an atomically flat interface when initial nitridation was employed.⁹ Hwang *et al.*¹⁰ reported that MOCVD-grown GaN films on prenitrided *c*-plane sapphire substrates gave rise to enhanced electrical and optical properties despite the rough surface morphology. Most recently, an atomic force microscopy (AFM) study of the nitrided sapphire substrates in MOCVD has been conducted.¹¹ This study shows that extended nitridation time with a fixed nitridation temperature (1050 °C) can result in a high density of nitrided protrusions on sapphire substrates, favoring three-dimensional growth of the successive GaN film. Also, nitridation of the sapphire substrate by nitrogen plasma in electron cyclotron resonance assisted molecular beam epitaxy (MBE)¹² has been reported to improve the quality of GaN films.

It is apparent from the studies mentioned above that nitridation of sapphire substrates can enhance the quality of III–V nitride layers grown by MOCVD. To our knowledge, no such study of the nitridation process and its effect on growth by the RMBE technique has yet been performed. Notably, the activation process of the nitrogen species and the growth surface mediated chemical reaction, which leads to high growth rates (~ 2 $\mu\text{m/h}$),¹³ are different in RMBE when compared with MOCVD. However, although there are a number of differences between MOCVD and RMBE, the nitridation process that involves active nitrogen species generated from ammonia molecules on a heated sapphire substrate is similar in both cases. Yamamoto¹⁴ observed in the MOCVD process that nitridation of sapphire surface occurs at a temperature higher than 800 °C. Incidentally, this is the

^{a)}Electronic mail: w-kim1@uiuc.edu

^{b)}On sabbatical leave at Wright Laboratory, Wright Patterson AFB, OH 45433 under a University Resident Research Program funded by AFOSR.

TABLE I. Experimental conditions and area roughness data for nitrided surfaces and AlN layers.

Sample	AlN growth temperature (°C)	Nitridation temperature (°C)	Nitridation time (min)	Thickness (nm)	Growth rate (nm/h)	R_A (nm)	R_{rms} (nm)
N0 _a	0.09	0.11
N1a	...	800	1	0.13	0.17
N5a	...	800	5	0.18	0.22
N15a	...	800	15	0.17	0.21
N30a	...	800	30	0.18	0.23
N1b	...	850	1	0.07	0.08
N5b	...	850	5	0.10	0.13
N15b	...	850	15	0.12	0.15
N1A40a	800	800	1	40	40	0.43	0.58
N10A40a	800	800	10	40	40	0.34	0.44
N10A120a	800	800	10	120	40	1.29	1.67
N10A130a	800	800	10	130	130	0.44	0.62
N1A40b	850	850	1	40	210	0.37	0.48
N1A100b	850	850	1	100	210	0.64	1.00
N1A210b	850	850	1	210	210	3.74	4.90

nitridation and growth temperature used in our RMBE process. Thus, in this article, we discuss an AFM and x-ray photoelectron spectroscopy (XPS) study aimed at determining the optimum conditions for nitridation and AlN growth, which provide smooth surfaces for the subsequent epitaxial growth of AlN buffer layers and GaN, respectively. This was accomplished by determining the effect of nitridation parameters on the surface roughness and chemical state of the nitrided surface, and on the surface roughness of AlN buffer layers grown after the nitridation process by RMBE.

II. EXPERIMENT

A conventional Riber 1000 MBE system was converted to a RMBE system in which ammonia gas was supplied through an ammonia injector. The temperature of the ammonia injector was maintained at 300 °C to avoid condensation of ammonia inside the injector tube during growth. The *c*-plane sapphire substrates were first degreased by dipping into warm trichloroethane, then cleaned with acetone and methanol, and finally rinsed with DI water. They were then etched in a hot solution of H₂SO₄:H₃PO₄=3:1 for 20 min and rinsed with DI water before being blow-dried with filtered nitrogen gas. Molten indium was used to mount the etched substrate on a silicon template held on a Mo block with carbon screws, and loaded into the sample introduction chamber. Before the initiation of growth, the substrates were heated to about 850 °C and held for about 3 min to desorb any surface contaminants. This was followed by nitridation, which was carried out by exposing the substrate to ammonia (16 sccm) while keeping the substrate temperature at 800 or 850 °C for a certain period of time. For some of the experiments, AlN was subsequently deposited. All of the growth parameters are listed in Table I. The thickness of the AlN layer was measured by an alpha-stepper and cross-sectional scanning electron microscopy. The surface morphology of the samples was investigated by AFM. To obtain reliable AFM images as well as surface roughness data, at least five

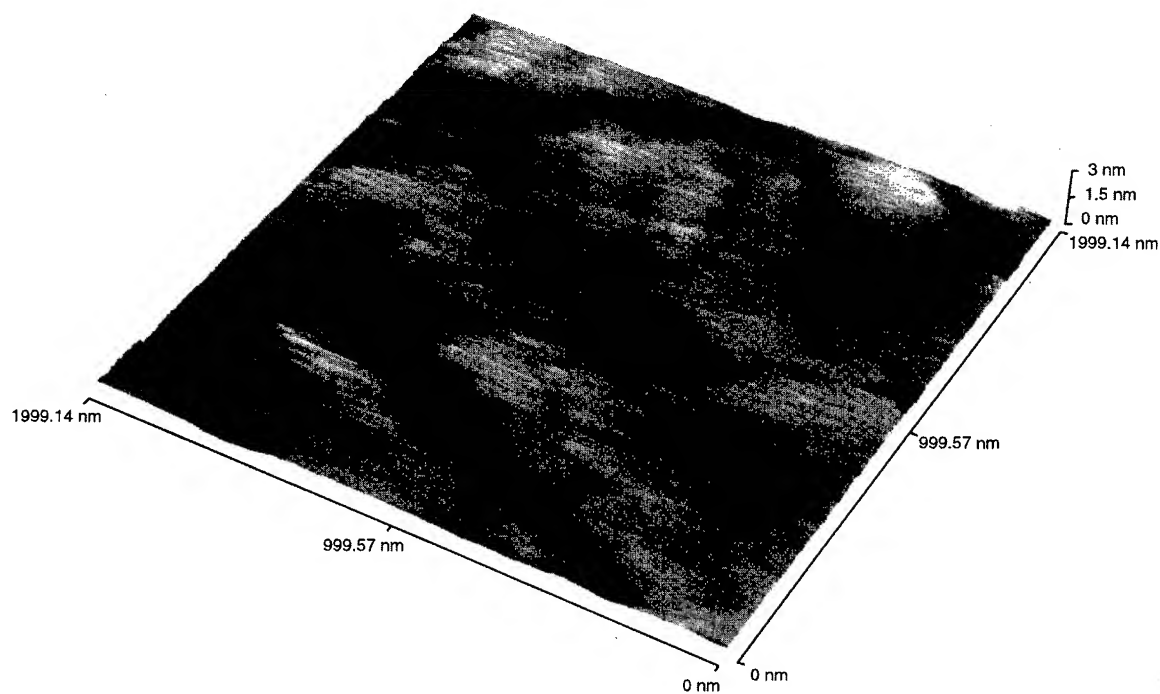
regions of each sample were examined. The arithmetic mean deviation R_A and root mean square deviation R_{rms} of AFM data values for a given area were used to describe the surface roughness of the samples. The nitrided substrates were analyzed using a PHI 5400 XPS with a Mg target to detect the incorporation of nitrogen species into the nitrided surface.

III. RESULTS AND DISCUSSION

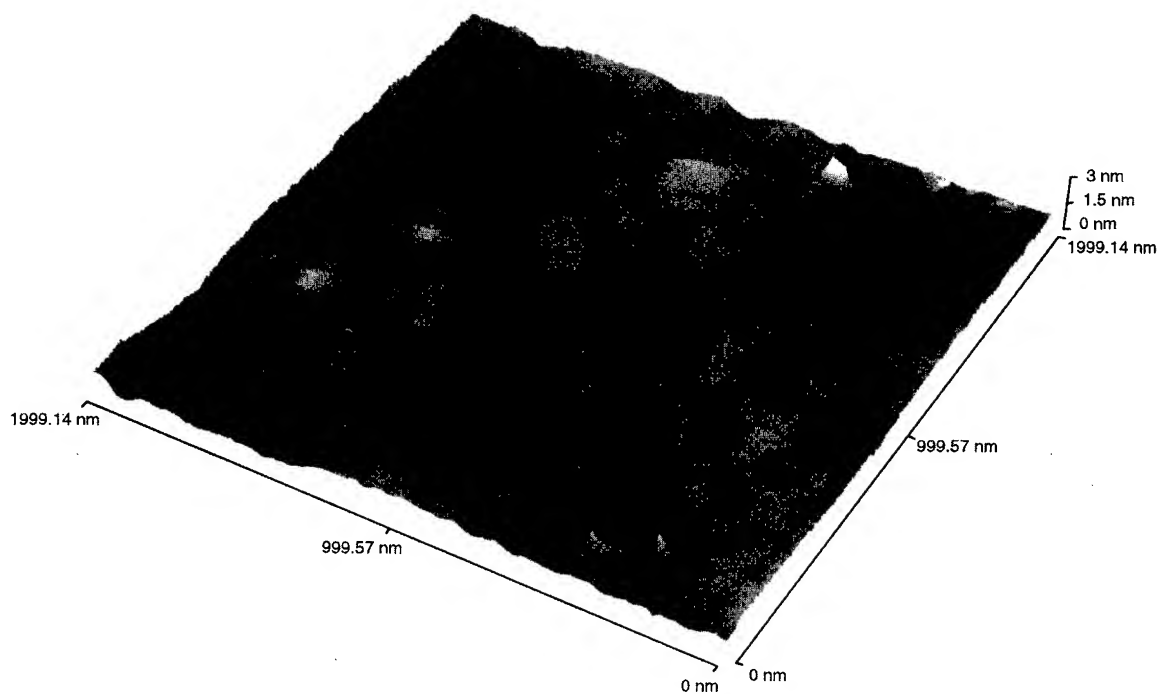
A. Nitridation

Figures 1(a) and 1(b) depict three-dimensional AFM images of the surfaces of the sapphire substrates after thermal desorption at 850 °C for 3 min and exposure to NH₃ gas at 800 °C for periods of 0 and 30 min (samples N0_a and N30a, respectively). Substrate surface damage is clearly visible on the surfaces of each of the samples due to the substrate polishing process. A low density ($\sim 10^8$ cm⁻²) of surface outgrowths was observed after 30 min of nitridation as shown in Fig. 1(b). The presence of surface damage does not appear to have influenced the formation of the protrusions and there is no clear correlation between the positions at which the protrusions have formed and the local surface topography. Uchida *et al.*¹¹ observed similar protrusions after 5 min of nitridation at 1050 °C in a MOCVD system, however, the density was approximately three orders of magnitude greater than that observed in Fig. 1(b). It is likely that the combination of a higher substrate temperature and background ammonia pressure promotes a more rapid nitridation reaction leading to a higher density of protrusions in MOCVD-grown samples.

To further characterize the influence of temperature on surface topography during nitridation, the substrate temperature was increased to 850 °C and fresh samples were exposed to ammonia flow (16 sccm) for periods of 1, 5, and 15 min (samples N1b, N5b, and N15b, respectively). Figure 2 shows the variation in surface roughness as a function of the



(a)



(b)

FIG. 1. AFM images of the surfaces of sapphire substrates after thermal desorption at 850 °C, and exposure to ammonia gas flow (16 sccm) for (a) 0 min (bare sapphire), (b) 30 min at a substrate temperature of 800 °C.

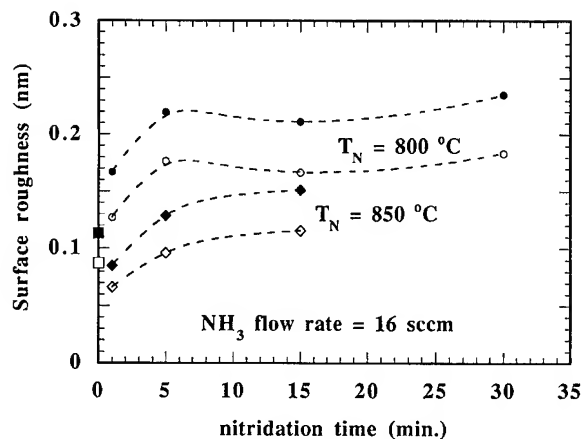
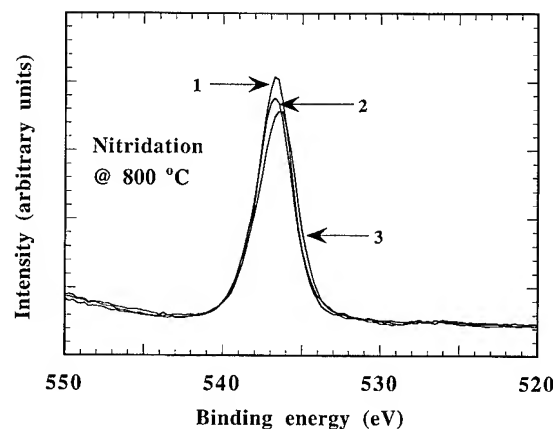


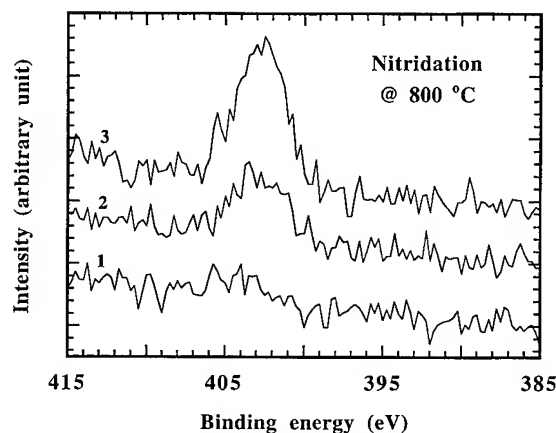
FIG. 2. Variation in surface roughness with nitridation duration for two different temperatures (800 and 850°C). Upper two lines correspond to nitridation at 800°C and lower two lines to 850°C . 0 min nitridation corresponds to bare sapphire substrate after thermal desorption at 850°C . Open markers are for R_A and closed ones for R_{rms} roughness. Dotted lines connecting data points are fitted curves. The error range for the data is $\pm 10\%$.

nitridation time, t_N , for the two different substrate temperatures T_N . The roughness data for the surfaces nitrided at 850°C are lower than for those nitrided at 800°C for the nitridation periods used in this investigation as shown in Fig. 2. A smoothing of the mechanical damage for the samples nitrided at 850°C was apparent (not shown in this article), contributing to the lower surface roughness. For the nitridation temperature of 850°C , the data suggest an initial decrease in surface roughness after 1 min nitridation. The magnitude of this decrease is comparable to the error in the roughness measurements, however, the form of the roughness curves as a function of nitridation time at the two temperatures are similar. Thus, the possibility of an initial decrease in roughness during nitridation at 850°C should not be ruled out since nitridation at this temperature does yield sapphire surfaces with lower roughness than those nitrided at 800°C .

The nitridation reaction is believed to yield $\text{AlN}_x\text{O}_{1-x}$ via the exchange of oxygen atoms from sapphire and nitrogen atoms from ammonia.¹¹ XPS analyses were carried out to confirm the incorporation of nitrogen atoms into the sapphire substrate during the nitridation process. Figure 3(a) and 3(b) show high resolution XPS scans for oxygen 1s and nitrogen 1s lines for a bare sapphire substrate (sample N0, curve 1), the sapphire substrate nitrided for 1 min at 800°C (sample N1a, curve 2), and the sapphire substrate nitrided for 30 min at 800°C (sample N30a, curve 3). There is essentially no nitrogen 1s peak for the bare sapphire substrate, as expected. Based on the development of the nitrogen 1s peak with nitridation time as shown in Fig. 3(b), it is speculated that nitrogen atoms are incorporated into the surface of the sapphire substrate even with only 1 min of nitridation. From the XPS scan after 30 min nitridation, it can be seen that the oxygen 1s line has been down shifted by about 0.25 eV relative to the spectra obtained from the bare sapphire sample and the sample nitrided for 1 min. Although further



(a)



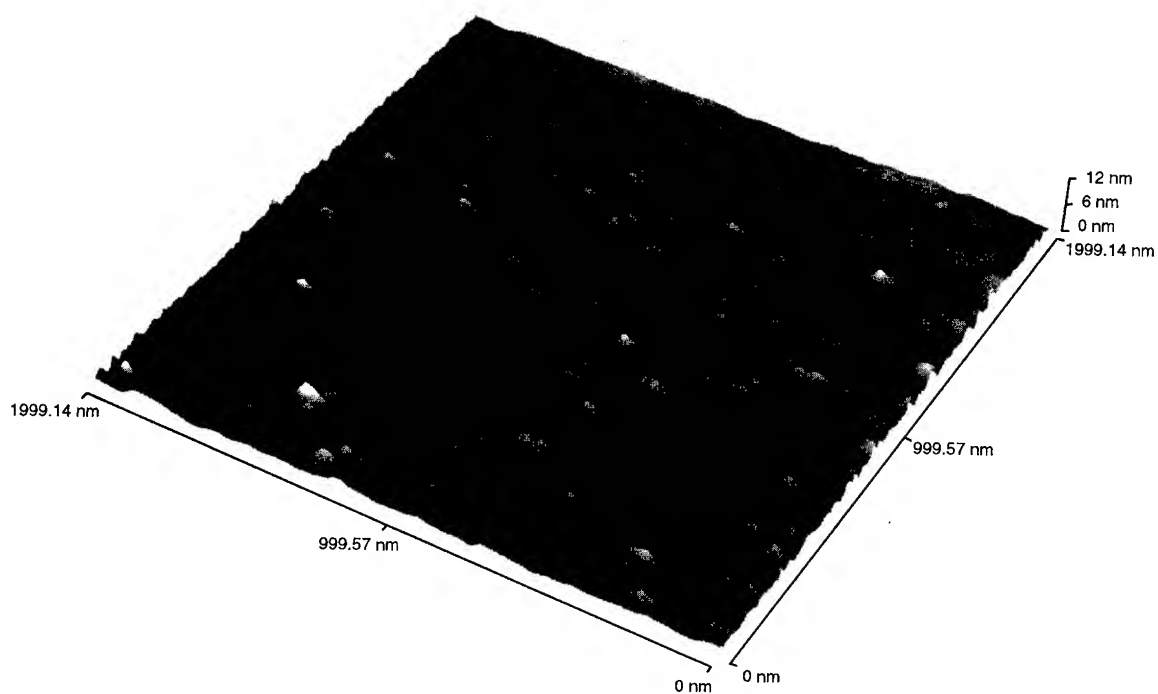
(b)

FIG. 3. High resolution XPS scan for (a) O 1s line, (b) N 1s line. In the figures, curve 1 corresponds to the bare sapphire substrate, curve 2 to the sapphire substrate that underwent 1 min of nitridation, and curve 3 to 30 min of nitridation.

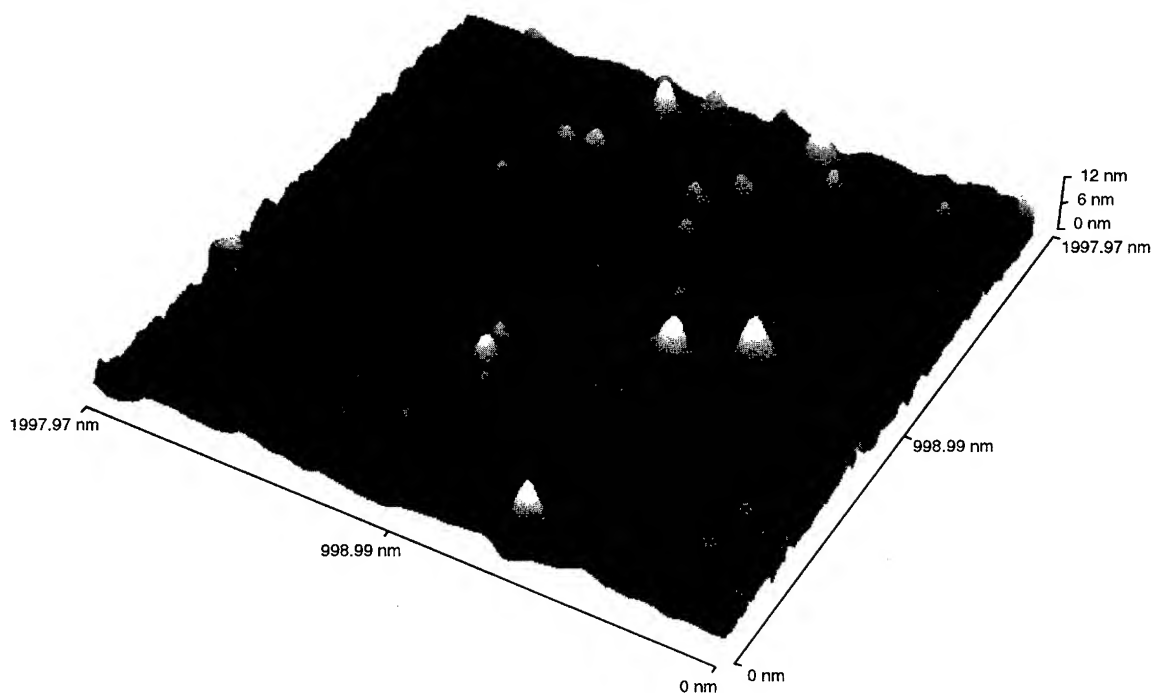
investigation is required, it is possible that this is due to a chemical shift arising from the generation of a significant number of new bonding states between oxygen and nitrogen atoms in the nitrided layer.

B. AlN growth

It has been reported⁵ that AlN buffer layers grown by MOCVD at higher substrate temperatures have a rougher surface morphology than those grown at lower temperatures ($\approx 600^\circ\text{C}$). It has also been reported that buried polycrystalline or amorphous AlN layers deposited by MOCVD at lower temperatures are converted into single crystal form during substrate heating and GaN deposition at elevated temperature ($\geq 1000^\circ\text{C}$),⁵ however, it is not clear how the conversion affects the growth of the GaN. Unlike the AlN buffer layers grown by MOCVD, RMBE-grown AlN buffer layers are not subjected to temperatures on the order of 1000°C since the growth temperature of our GaN layers is the same, or lower than, the growth temperature of the AlN buffer. The surface roughness and crystalline quality of the as-grown AlN buffer layer will therefore have a more direct influence

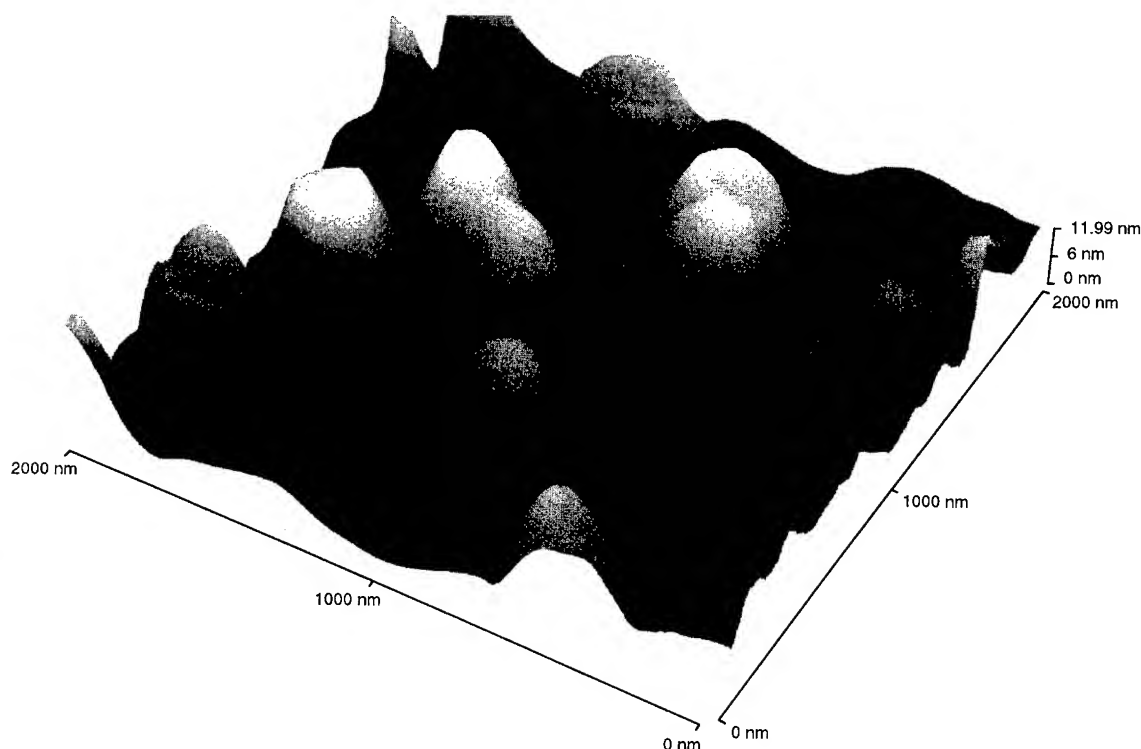


(a)



(b)

FIG. 4. AFM images for AIN surface grown with growth rate of 210 nm/h up to various thicknesses. The conditions for the growth and the nitridation are detailed in Table I. (a) Sample N1A40b, (b) sample N1A100b, (c) sample N1A210b.



(c)

FIG. 4. (Continued).

on the quality of the GaN layer deposited on it. Table I summarizes the experimental conditions for the nitridation processes we investigated. It also lists the growth conditions for the AlN buffer, layers grown on various nitrided surfaces described earlier in the text.

In order to understand the effect of nitridation time on the roughness of the AlN layers, two 40 nm-thick AlN films (samples N1A40a and N10A40a) were grown with 1 and 10 min nitridation, respectively, keeping all other growth conditions identical. These periods were chosen since the former represents a relatively smooth nitrided surface and the latter a rougher surface as discussed in the previous section. From Table I it can be seen that, although the substrate surface is more rough after 10 min nitridation compared with only 1 min of nitridation, the surface of the AlN layer with 10 min nitridation is smoother. The surfaces of the AlN layers are characterized by a high density of protrusions, with average diameters of 100 and 50 nm for the 1 min and 10 min nitridation times, respectively. The quoted feature dimensions should be treated with some caution, however, as they represent upper bounds on the feature sizes since the images obtained using AFM actually represent a convolution of the tip and surface geometries. Although the latter film has a lower R_{rms} surface roughness, there is a higher density of potential nucleation sites for subsequent GaN growth. It is not yet clear how this will affect the quality of subsequent GaN layers. In an attempt to obtain smoother AlN surfaces with large AlN crystallites, thicker (120 nm) layers were

grown, one at the same growth rate as the previous layers (40 nm/h), and one at the higher growth rate of ~ 120 nm/h (samples N10A120a and N10A130a, respectively). Nitridation of the substrate was performed for 10 min in both cases.

An increase in film thickness at the same growth rate led to a significant increase in the coarseness of the AlN surface (the surface roughness increased by a factor of 4 for sample N10A120a), with the crystallites having an average diameter of approximately 143 nm. The sample grown at the higher growth rate (N10A130a), however, having almost the same thickness as sample N10A120a, was found to be significantly smoother. The main contributions to the surface roughness of this sample are isolated surface protrusions.

Based on the data from the previous section, it is evident that nitridation at 850 °C for 1 min results in the smoothest substrate surface. Using this condition, three AlN layers (N1A40b, N1A100b, and N1A210b) were grown at the higher growth rate of 210 nm/h up to three different thickness (40, 100, and 210 nm, respectively). Figures 4(a)–4(c) show the AFM images of these three samples. The surface roughness increases with thickness as shown in Fig. 5. An inspection of Figs. 4(a)–4(c) indicates that substantial coarsening of the AlN surface has occurred at a thickness of 100 nm [Fig. 4(b)], with further surface roughening occurring at greater thickness [Fig. 4(c)]. It should again be noted that the images presented in Fig. 4 represent a convolution of tip and surface geometries; while the trend towards substantial sur-

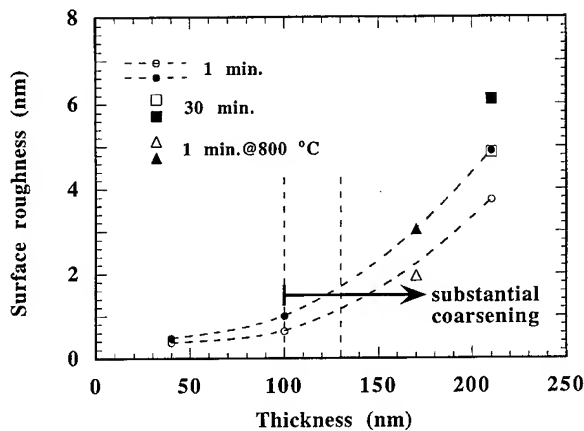


FIG. 5. Variation in surface roughness of AlN layer with thickness. Nitridation was carried out at 850 °C except one data point with thickness of 170 nm, which was nitrided at 800 °C. Open markers are for R_A and closed ones for R_{rms} roughness. Dotted lines connecting data points are fitted curves. The error range for the data is $\pm 10\%$.

face coarsening is clear, the protrusions may in fact be sharper than their representations in the figure.

It is evident from our investigation that, over the ranges of conditions investigated, nitridation period and growth temperature alone are not the primary factors determining the surface roughness of AlN layers. The most dramatic influences on the surface roughness of AlN are layer thickness and growth rate; for a given growth rate there is a significant coarsening of the surface morphology with increasing thickness suggesting a microstructural ripening mechanism which is not critically dependent on the initial density of nucleation sites provided by the nitrided surface. Our study indicates therefore that the optimum growth conditions required to obtain the smoothest AlN buffer layers are found with film thickness in the range 40–130 nm deposited at growth rates of between 130 and 210 nm/h, since lower growth rates increase the density of surface protrusions.

IV. CONCLUSION

The effect of different nitridation conditions on the surface roughness of sapphire substrates and AlN layers subsequently deposited has been studied using AFM. At the nitridation temperatures used in this report (800 and 850 °C), the surface roughness increases with nitridation duration up to 5

min and then saturates with further nitridation. Within the scope of the experiment, the roughness values for higher nitridation temperature are smaller overall due to a smoothing out of substrate surface damage. The incorporation of nitrogen species during the nitridation period has been confirmed using XPS.

The surface morphology of AlN layers deposited on nitrided surfaces in our system is influenced primarily by the layer thickness and growth rate. With increasing film thickness the surface undergoes a ripening transition leading to a rough, coarse-grained morphology. For a given layer thickness, an increase in the growth rate from 40 to 120 nm/h was found to give a reduction in the surface roughness by up to a factor of 3.

ACKNOWLEDGMENTS

The research is supported by the Office of Naval Research under Grant No. N00014-95-1-0324 monitored by Dr. Richard Brandt.

- ¹S. N. Mohammad, A. Salvador, and H. Morkoç, *Proc. IEEE* **83**, 1306 (1995).
- ²S. Nakamura, M. Senoh, N. Iwasa, and S.-I. Nagahama, *Jpn. J. Appl. Phys.* **34**, L797 (1995).
- ³S. Nakamura, M. Senoh, S.-I. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto, *Jpn. J. Appl. Phys.* **35**, L76 (1996).
- ⁴Z.-F. Fan, S. N. Mohammad, O. Aktas, A.E. Botchkarev, A. Salvador, and H. Morkoç, *Appl. Phys. Lett.* **69**, 1229 (1996); S. N. Mohammad, Z.-F. Fan, A. Salvador, O. Aktas, A.E. Botchkarev, W. Kim, and H. Morkoç, *ibid.* **69**, 1420 (1996); O. Aktas, W. Kim, Z.-F. Fan, A.E. Botchkarev, A. Salvador, S.N. Mohammad, B. Sverdlov, and H. Morkoç, *Electron. Lett.* **31**, 1389 (1995).
- ⁵I. Akasaki, H. Amano, Y. Koide, K. Hiramatsu, and N. Sawaki, *J. Cryst. Growth* **98**, 209 (1989); J. N. Kuznia, M. A. Khan, D. T. Olson, R. Kaplan, and J. Freitas, *J. Appl. Phys.* **73**, 4700 (1993).
- ⁶S. Nakamura, *Jpn. J. Appl. Phys.* **30**, L1705 (1991).
- ⁷Y. Koide, N. Itoh, K. Itoh, N. Sawaki, and I. Akasaki, *Jpn. J. Appl. Phys.* **27**, 1156 (1988).
- ⁸H. Kawakami, K. Sakurai, K. Tsubouchi, and N. Mikoshiba, *Jpn. J. Appl. Phys.* **27**, L161 (1988).
- ⁹K. Masu, Y. Nakamura, T. Yamazaki, T. Shibata, M. Takahashi, and K. Tsubouchi, *Jpn. J. Appl. Phys.* **34**, L760 (1995).
- ¹⁰C.-Y. Hwang, M. J. Schurman, W.E. Mayo, Y. Li, Y. Lu, H. Liu, T. Salagaj, and R. A. Stall, *J. Vac. Sci. Technol. A* **13**, 672 (1995).
- ¹¹K. Uchida, A. Watanabe, F. Yano, M. Koguchi, T. Tanaka, and S. Minagawa, *J. Appl. Phys.* **79**, 3487 (1996).
- ¹²T.D. Moustakas, T. Lei, and R. J. Molnar, *Physica B* **185**, 36 (1993).
- ¹³W. Kim, O. Aktas, A.E. Botchkarev, A. Salvador, S. N. Mohammad, and H. Morkoç, *J. Appl. Phys.* **79**, 7657 (1996).
- ¹⁴A. Yamamoto, M. Tsujino, M. Ohkubo, and A. Hashimoto, *J. Cryst. Growth* **137**, 415 (1994).

Influences of the [(Ba,Sr)TiO₃]-modified RuO₂ interface on the dielectric constant and current-voltage characteristics

Min-Seok Jeon^{a)} and Duck-Kyun Choi^{b)}

Department of Inorganic Materials Engineering, Hanyang University, 17 Haengdang-dong, Seongdong-ku, Seoul 133-791, Korea

(Received 5 February 1997; accepted 16 May 1997)

The electrical properties of (Ba,Sr)TiO₃ (BST) thin films on modified RuO₂ electrodes were investigated. The surfaces of rf magnetron sputtered RuO₂ electrode were modified by rf plasma etching, O₂ annealing, and overlayer coating of Pt. BST thin films were deposited on the various electrodes at identical conditions. There was no significant influence of the surface roughness on the phase formation and columnar, hillock-free, and dense BST films were obtained in all cases. BST films on different electrodes showed analogous dielectric constant when the effective electrode area was considered in calculation. But the roughened surface having large charge accumulation area is advantageous for carrying larger charge storage capacity. The leakage current of the BST film on RuO₂ electrode without any modification at 1.5 V was 5.06×10^{-5} and it increased to 7.92×10^{-5} and 1.07×10^{-4} A/cm² for the BST films on annealed RuO₂ and rf plasma etched RuO₂, respectively. The BST film on Pt/RuO₂/SiO₂/Si showed the lowest leakage current of 9.64×10^{-7} A/cm². The Schottky barrier height acquired from the temperature dependence of $I-V$ characteristics turned out to be 0.94, 0.79, 0.72, and 1.65 eV for the interface between the BST film and the RuO₂ reference, rf plasma etched RuO₂, annealed RuO₂, and Pt(50 nm)/RuO₂. © 1997 American Vacuum Society. [S0734-211X(97)14004-5]

I. INTRODUCTION

(Ba,Sr)TiO₃(BST) thin film is an attractive material for the application in high density dynamic random access memories (DRAMs)¹⁻⁴ because of the high relative dielectric constant and small variation in dielectric properties with frequency. In addition, BST thin film can exist in the form of paraelectric phase at room temperature by controlling the composition. Therefore, there is no fatigue and aging issues which are a major concern of Pb(Zr,Ti)O₃ or (Pb,Ln)(Zr,Ti)O₃.

BST films for DRAMs require a new electrode material due to the significant degradation of dielectric properties by the formation of low dielectric constant interfacial layer (SiO₂) in using conventional highly doped poly-Si. Noble metal electrodes like Pt have been adopted as storage electrodes because of high resistance to oxidation and stability, while a suitable etching technology for micropatterning of Pt is not fully developed. Fortunately, RuO₂ can be used as a bottom electrode for DRAM capacitors because it has low resistivity ($<100 \mu\Omega$ cm at room temperature) and a capability of blocking diffusion.⁵⁻⁷ It can also be easily patterned by oxygen-chlorine gas or oxygen-fluorine gas plasma.⁸ However, the leakage current of BST thin films is larger on this oxide electrode than on Pt. Moreover, RuO₂ has a rougher surface than Pt. Such roughness is originated from the faceted grains of RuO₂. The leakage current mechanism of BST films on Pt has been reported in terms of thermally activated electron emission process by a few authors.⁹⁻¹² However, there are few reports about the investigation in the

leakage current of BST films on RuO₂ and the influence of the surface roughness of RuO₂ on the electrical properties of BST films.

In this study, a RuO₂ bottom electrode was modified intentionally by oxygen-argon gas plasma etching, thermal etching (annealing), and overlayer Pt film coating to study the role of the interface between BST and bottom electrodes on the electrical properties. The leakage current of BST/RuO₂ and BST/Pt/RuO₂ is analyzed by comparing the Schottky barrier height.

II. EXPERIMENT

RuO₂ films (200 nm) were deposited by rf magnetron sputtering on thermally grown SiO₂ (200 nm)/Si substrates at a substrate temperature of 300 °C and gas pressure of 1×10^{-2} Torr. Plasma etching with Ar+O₂ and thermal etching (annealing) at 700 °C in atmospheric oxygen were adopted to modify the surface of as-deposited RuO₂ electrode (RuO₂ reference). For preparing hybrid electrode, an overlayer Pt film of 50 nm was deposited on the RuO₂ reference by dc magnetron sputtering at room temperature and a gas pressure of 3.5×10^{-3} Torr. (Ba,Sr)TiO₃ films (200 nm) were deposited at identical conditions by rf magnetron sputtering on each type of RuO₂ electrode with a (Ba_{0.5}Sr_{0.5})TiO₃ ceramic target at a substrate temperature of 600 °C and gas pressure of 1×10^{-2} Torr. The sputtering conditions are summarized in Table I.

The root mean square (rms) roughness of several electrodes was measured by atomic force microscopy (AFM) in the scanning range of $0.6 \mu\text{m} \times 0.6 \mu\text{m}$. The phase formation and crystallinity of the (Ba,Sr)TiO₃ films were investigated by x-ray diffraction (XRD) with Cu K α radiation (40 kV, 30 mA). The scanning range was from 20° to 50° and the scan

^{a)}Electronic mail: msjeon@hmail.hanyang.ac.kr

^{b)}Electronic mail: duck@email.hanyang.ac.kr

TABLE I. Sputtering conditions of RuO₂ and Ba(Sr,Ti)O₃ thin films.

	Deposition conditions	
	RuO ₂ film	(Ba,Sr)TiO ₃ film
Target	Cold pressed powder	(Ba _{0.5} Sr _{0.5})TiO ₃ sintered ceramics
Substrate temperature	300 °C	600 °C
Radio frequency power density	2.22 W/cm ²	16.035 W/cm ²
Sputtering gas	10% O ₂ -90% Ar	20% O ₂ -80% Ar
Gas pressure	1 × 10 ⁻² Torr	1 × 10 ⁻² Torr
Film thickness	200 nm	200 nm

speed was 7°/min. The surface morphology of the BST films and modified RuO₂ electrodes was analyzed by scanning electron microscopy (SEM). For electrical measurements, Al/BST/modified RuO₂/SiO₂/Si capacitors (MIM configuration) were fabricated. Al (300 nm in thickness) dots for top electrodes were thermally evaporated and patterned using a metal shadow mask of 200 μm or 400 μm in diameter. The dielectric constant of the (Ba,Sr)TiO₃ films was calculated from the capacitance measured by a HP 4285A precision LCR meter at a frequency of 100 kHz. The leakage current of the BST films was measured by a HP 4145B semiconductor parameter analyzer. *I*-*V* measurement was carried out with a staircase voltage ramp of 0.025 V which was applied to the top electrode with the bottom electrode grounded. To investigate the dependence of leakage characteristics on the polarity of bias voltage, leakage current was analyzed from -10 to 10 V at a step of 0.05 V. The temperature dependent

I-*V* characteristics of the BST thin films were obtained at different temperatures in the range of 310-380 K.

III. RESULTS AND DISCUSSION

The surface morphology of the bottom electrodes is given in Fig. 1. An RuO₂ electrode prepared at standard process (RuO₂ reference) was shown to have a rough and faceted surface. A rougher surface than the RuO₂ reference was obtained by rf plasma etching [Fig. 1(b)]. Pt (50 nm)/RuO₂ substrate was observed to have hillock-free and similar surface morphology to the RuO₂ reference. There is no significant difference of the surface morphology between the RuO₂ reference and annealed RuO₂ from SEM micrographs. A closer look of the surfaces was done by AFM. Thermally etched RuO₂ reveals that the surface roughness increased in comparison to the RuO₂ reference (Fig. 2). The rms roughnesses of the substrates were 4.4, 9.5, 8.6, and 4.5 nm for the RuO₂ reference, rf plasma etched RuO₂, annealed RuO₂, and Pt/RuO₂, respectively. An increase in the roughness of the annealed RuO₂ may be attributed to the volatile phase formation of RuO₃ or RuO₄ (Ref. 13) by the reaction between RuO₂ and O₂ at the grain boundaries at 700 °C.

X-ray diffraction patterns of the BST films deposited on the modified RuO₂ electrodes having different surface roughness are shown in Fig. 3. XRD peaks from Pt and Si substrates are also shown. When the peaks only from the BST films are considered, those are quite alike. Especially, the major BST (110) peak at about 32° does not show a big difference. This result indicates that there is no significant influence of the surface morphology of the bottom electrodes on the phase formation of the BST films.

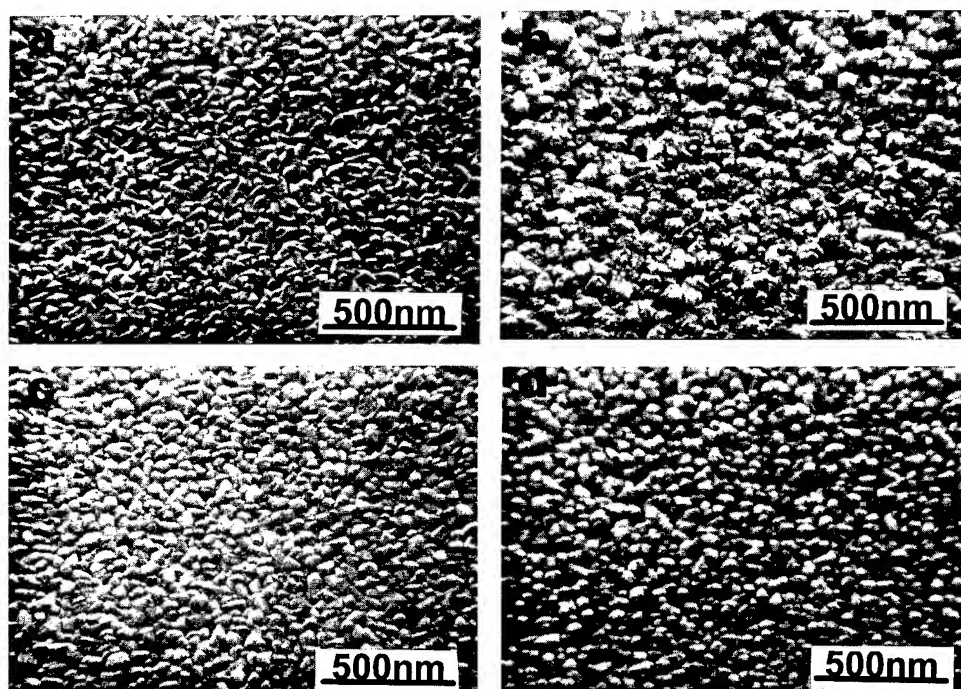


Fig. 1. SEM micrographs of bottom electrodes; (a) RuO₂ reference, (b) rf plasma etched RuO₂, (c) thermally etched RuO₂, and (d) Pt/RuO₂.

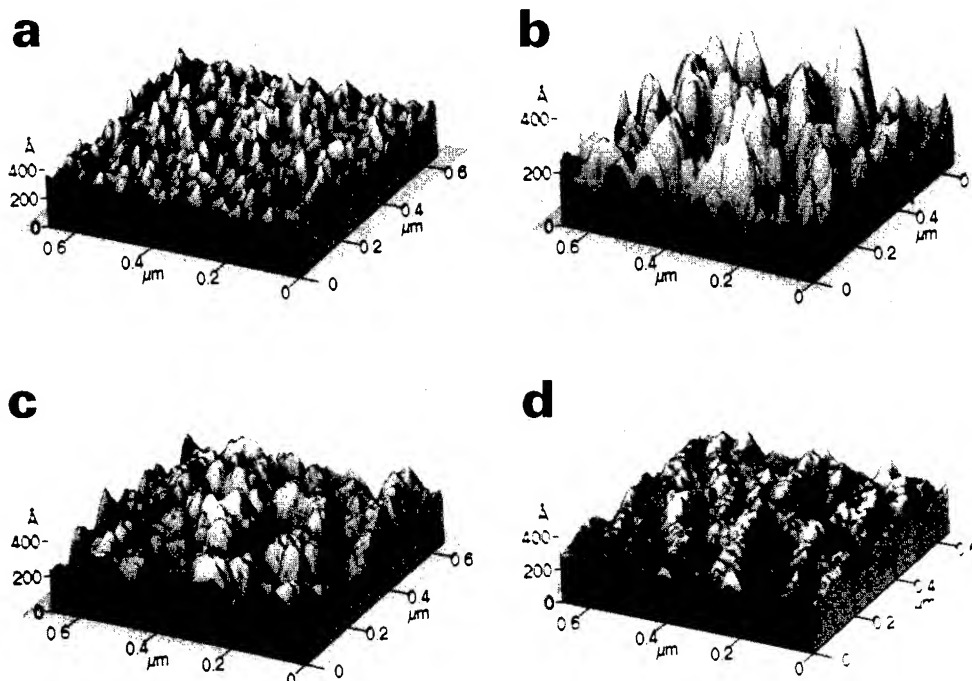


FIG. 2. AFM micrographs of bottom electrodes; (a) RuO₂ reference, (b) rf plasma etched RuO₂, (c) thermally etched RuO₂, and (d) Pt/RuO₂.

Figure 4 shows SEM micrographs of BST films. There is no remarkable difference of the surface morphology between the BST films regardless of the bottom electrodes.

A. Dielectric properties

Figure 5 illustrates the dielectric constant and dissipation factor ($\tan \delta$) of BST film capacitors. It is shown that the BST film on the rf plasma etched RuO₂ has the highest apparent dielectric constant among all capacitors and the apparent dielectric constants are in good agreement with the surface roughness of the bottom electrodes. Such a difference in

the apparent dielectric constants may not have originated from the inherent properties of the BST films because these films consist of almost identical grains and crystalline phase. Therefore, the effective electrode area due to the surface roughness may be responsible for such difference. To figure out the quasitru dielectric constant considering the effective area of the capacitor, each faceted tip of grains was assumed as a circular cone. The height and diameter of the cone were drawn from the rms roughness and grain size of RuO₂ (~50 nm), respectively. This approach was done for illustrating the influence of the roughened surface on the dielectric constant under the judgment that the rms roughness can be used for the approximate comparison between the specimens. The effective electrode area was calculated from the surface area of each circular cone and the number of cones in the electrode area defined by a designed metal shadow mask ($1.257 \times 10^{-3} \text{ cm}^2$). The effective capacitor areas of four different electrodes considering the surface roughness were 2.552×10^{-3} , 2.688×10^{-3} , 2.658×10^{-3} , and $2.554 \times 10^{-3} \text{ cm}^2$ in order of the modification. The apparent dielectric constants were corrected by these numbers and also given in Fig. 5 as the quasitru dielectric constants (solid circle). Through the correction for the effective electrode area, the dielectric constant changed remarkably and showed no big difference between the BST films on the various electrodes. This surface roughness, however, would result in the improvement of charge storage capacity per unit cell in real application. The dissipation factor of the BST films crystallized on the various electrodes remained between 0.015 and 0.025.

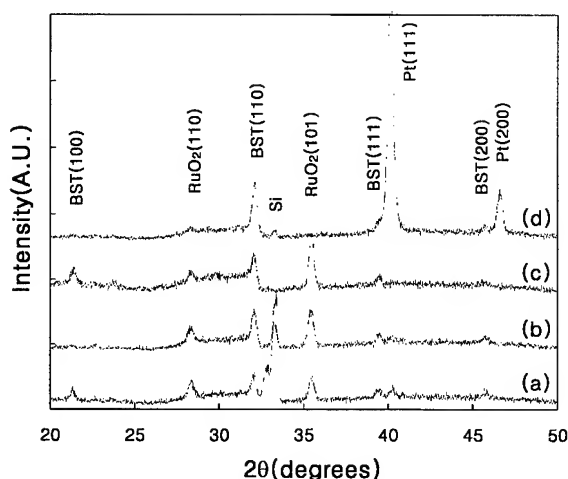


FIG. 3. X-ray diffraction patterns of BST thin films deposited on (a) RuO₂ reference, (b) rf plasma etched RuO₂, (c) thermally etched RuO₂, and (d) Pt/RuO₂.

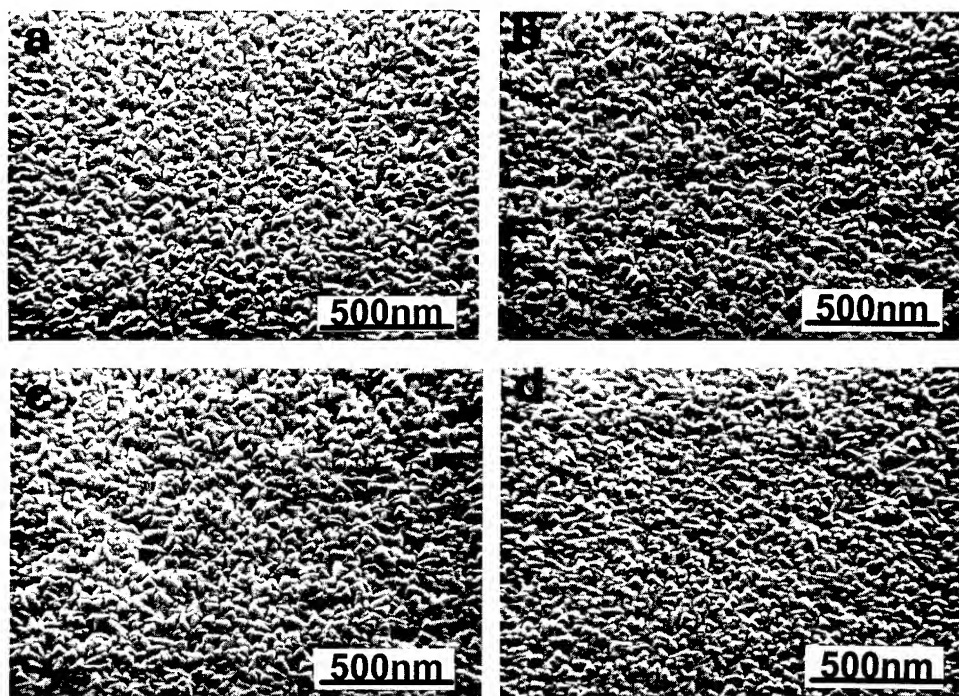


FIG. 4. SEM micrographs of BST films deposited on (a) RuO₂ reference, (b) rf plasma etched RuO₂, (c) thermally etched RuO₂, and (d) Pt/RuO₂.

B. Current-voltage characteristics

Figure 6 shows the leakage current characteristics of BST thin film capacitors with the polarity of applied voltage. The BST/Pt/RuO₂ capacitor has a leakage current of 9.64×10^{-7} A/cm² at 1.5 V and shows high asymmetry. This result implies that the leakage current is controlled by Schottky (interface-limited) emission. Such asymmetry is from the large work function difference between BST (~ 4.0 eV)¹⁴ and P (5.65 eV).¹⁵ Poole-Frenkel (bulk-limited) emission, on the other hand, is related to the barrier lowering of trap sites in BST films by applied field. Therefore, there is no

polarity dependence of Poole-Frenkel emission on I - V characteristics. The BST film on RuO₂ reference does not show such asymmetry in the I - V curve and shows larger leakage currents than the BST film on Pt/RuO₂. Such an increase in the symmetry of the I - V curve indicates that bulk-limited emission is somewhat effective relative to interface-limited emission which is favorable for the BST/Pt/RuO₂ capacitors. For the BST film capacitors prepared on the plasma etched and annealed RuO₂ electrodes,

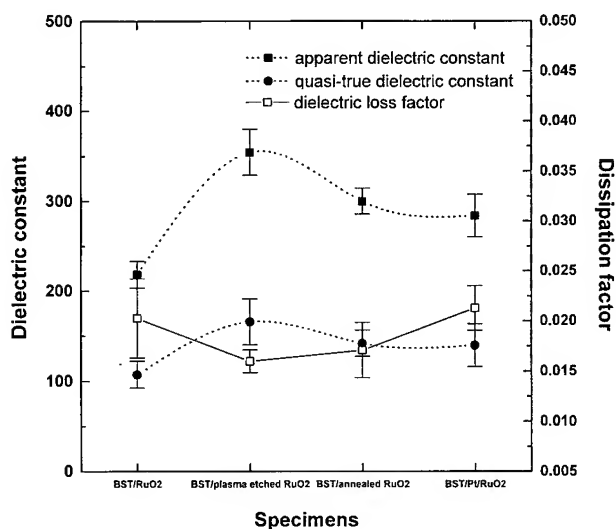


FIG. 5. Dielectric constants of BST films for various bottom electrodes. (Dotted line is inserted for representing the range of dielectric constants.)

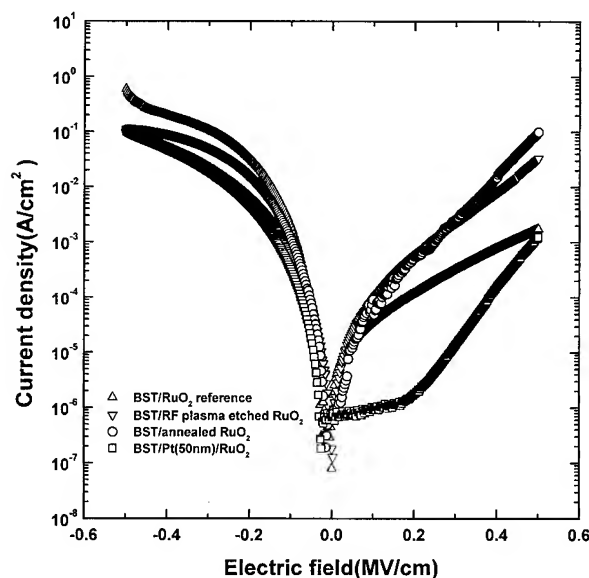


FIG. 6. Leakage current vs applied voltage of BST thin film capacitors on modified RuO₂ electrodes.

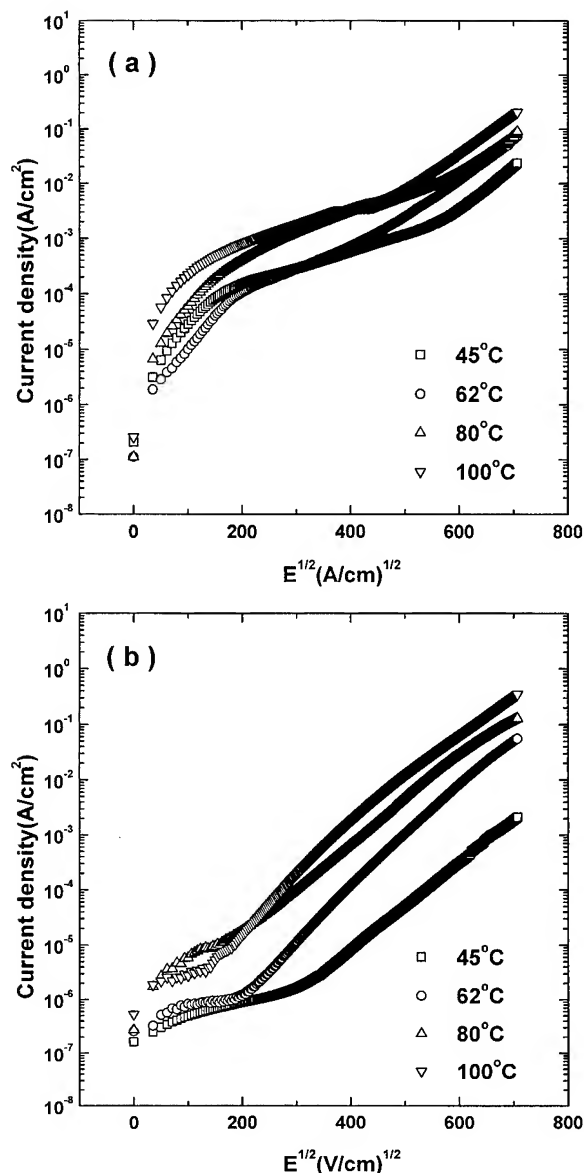


Fig. 7. J vs $E^{1/2}$ characteristics of BST thin films deposited on (a) RuO₂ reference and (b) Pt/RuO₂.

more symmetric I – V characteristics were obtained. The values of leakage current at 1.5 V are 5.06×10^{-5} , 1.07×10^{-4} , and 7.92×10^{-5} A/cm² for the BST/RuO₂ reference, BST/RF plasma etched RuO₂, and BST thermally etched RuO₂, respectively. This increase in leakage current with the increase of the surface roughness may be ascribed to a local increase of field at the region of sharp tips on the bottom electrode. Therefore, after electrons are injected from the bottom electrode to the BST film over interfacial potential (Schottky barrier), they can be transported from columnar BST film to top electrode more easily through the high field region.

Figure 7 shows J vs $E^{1/2}$ characteristics of the BST thin films on the RuO₂ reference and Pt/RuO₂ as a function of measurement temperature. The leakage current increases remarkably as the temperature is elevated. The linear region in

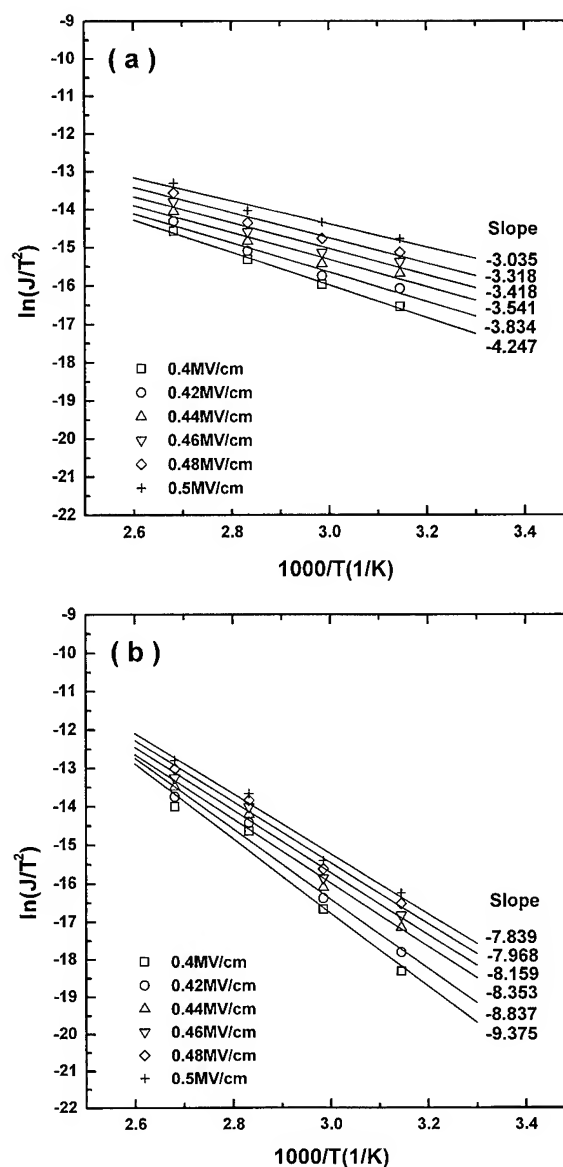


Fig. 8. $\ln(J/T^2)$ vs $1000/T$ plots of (a) BST/RuO₂ reference and (b) BST/Pt/RuO₂.

the J vs $E^{1/2}$ curve at high field can correspond to either Schottky emission at the interface between the BST and bottom electrodes or Poole–Frenkel emission within the BST thin films. Therefore, there is a mixed contribution of the two mechanisms to the leakage characteristics of the BST thin films. The dominant operating mechanism, however, could be distinguished from the I – V curve with the polarity of applied bias, as explained in Fig. 6. In this study, the leakage characteristics of the BST films were mainly investigated in the aspect of Schottky (interface-limited) emission under the assumption that there might be no big difference in the bulk of BST films as a result of the identical deposition conditions.

The temperature dependence of leakage current was further studied at different electric fields in the linear region of the J vs $E^{1/2}$ curve. In Fig. 8, the logarithm of J/T^2 is plotted as a function of reciprocal temperature for the BST thin films

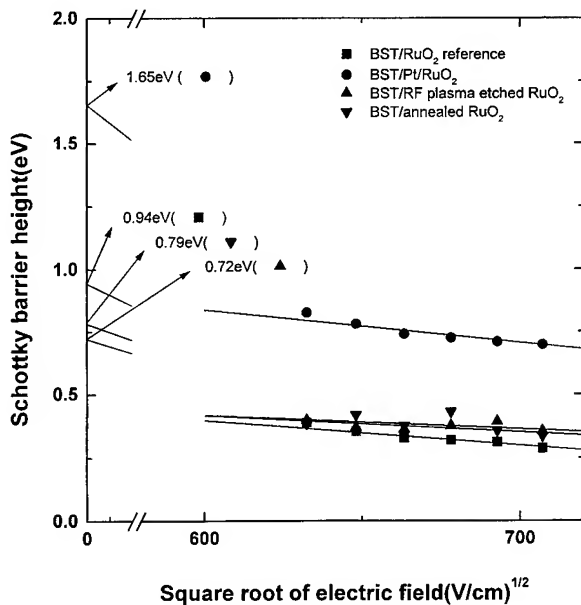


FIG. 9. Extrapolation of Schottky barrier height for the interface between BST films and modified RuO₂ bottom electrodes.

on the RuO₂ reference and Pt/RuO₂ electrodes. The Schottky barrier height including Schottky effect can be deduced from the equation of

$$q\Phi_B = k[-1000S^* + (qE/4\pi\epsilon_0\epsilon_r k^2)^{1/2}],$$

where S^* represents the slope of $\ln(J/T^2)$ vs $1000/T$ curve at each electric field (E) and ϵ_r is the relative dielectric constant. The quasitru dielectric constant in Fig. 5 was used for ϵ_r .

The Schottky barrier height of the interface between the BST and modified RuO₂ electrodes at $E=0$ is extrapolated from each $q\Phi_B$ at the linear region of electric field (0.4–0.5 MV/cm) (Fig. 9). Recently, Y. Shimada *et al.*¹⁰ reported that Schottky barrier height between metal-organic chemical-vapor deposition BST and Pt is about 0.8 eV. And W. H. Hsu *et al.*¹¹ obtained a Schottky barrier height of 0.86 eV from sputtered (Ba_{0.7}Sr_{0.3})TiO₃ and e-beam deposited Pt. It is known that practical Schottky barrier height can be influenced by defects and trap sites, carrier density, or interface states.¹⁶ These inhomogeneous states can be easily affected by deposition methods, processing conditions, or annealing conditions. Hwang *et al.*¹² reported a Schottky barrier height of 1.5–1.6 eV for sputtered (Ba_{0.5}Sr_{0.5})TiO₃ film on dc sputtered Pt and this value is close to our value of 1.65 eV. The interfaces between the BST and RuO₂ based electrodes resulted in smaller values of the Schottky barrier height than the BST/Pt (50 nm)/RuO₂. These values were 0.94, 0.79, and 0.72 eV for the BST films on RuO₂ reference, rf plasma etched RuO₂, and annealed RuO₂, respectively. M. Tomkiewicz *et al.*¹⁷ reported that the work function of RuO₂ is about 4.9 eV. Since this value is 0.75 eV lower than that of Pt, the degree of rectification in the BST/RuO₂ interface is much less effective than the BST/Pt interface.

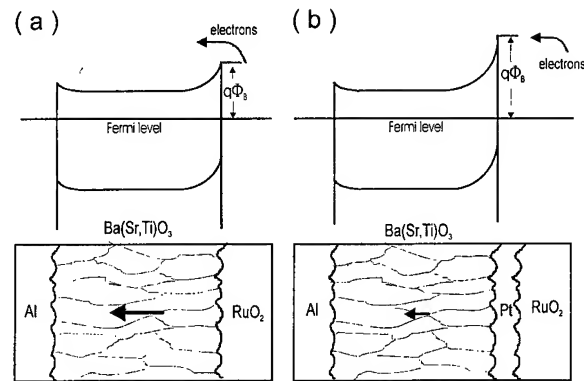


FIG. 10. Schematic band diagrams of (a) Al/BST/RuO₂ and (b) Al/BST/Pt/RuO₂.

Schematic band diagrams of the BST/RuO₂ and BST/Pt (50 nm)/RuO₂ are illustrated in Fig. 10. Very small blocking or actually ohmic contact may be formed in the Al/BST interface because there is almost no difference of the work function between Al (4.28 eV)¹⁵ and BST (~4.0 eV). After BST film contacts with RuO₂ or Pt/RuO₂, electrons will flow from BST to bottom electrode to equilibrate the Fermi levels. As a result, the conduction and valence band bend upward. Therefore, rectifying contact is formed in the BST/RuO₂ and BST/Pt/RuO₂ interface as shown in the diagram. From the above observation, the leakage current is more affected by the Schottky barrier height between the BST film and electrode rather than the surface roughness of electrode.

IV. CONCLUSIONS

Paraelectric (Ba,Sr)TiO₃ films were deposited on RuO₂ electrodes by rf magnetron sputtering. To investigate the dependence of electrical properties on the interface between BST and bottom electrode, as-deposited RuO₂ was modified by the methods of rf plasma etching, annealing (thermal etching), and overlayer Pt film coating. In all cases, the crystalline BST phase was obtained and all the BST films end up with columnar and dense grains with similar size. The quasitru dielectric constant calculated from the actual surface area considering the surface roughness was different from the apparent dielectric constant calculated simply from the designed electrode area. The charge storage capacity of BST film capacitors per unit area could be increased by roughening the surface. The leakage current increased as the surface roughness of RuO₂ electrode increased. This result may be due to the field concentration at the sharp point of the ferroelectric films. However, the effect of the surface roughness on the leakage current is much less than that of the Schottky barrier height due to the work function difference. The large difference in the work function between BST thin film and Pt makes the leakage current suppressed by decreasing the electron flow at the interface. In this study, the Schottky barrier height in the BST/RuO₂ and BST/Pt/RuO₂ turned out to be 0.94 and 1.65 eV, respectively.

- ¹K. Takemura, T. Sakuma, and Y. Miyasaka, *Appl. Phys. Lett.* **64**, 2967 (1994).
- ²T. Kuroiwa, *Jpn. J. Appl. Phys.* **33**, 5187 (1994).
- ³K. Takemura, *Jpn. J. Appl. Phys.* **34**, 5224 (1995).
- ⁴C. S. Hwang, *Appl. Phys. Lett.* **67**, 2819 (1995).
- ⁵R. G. Vadimsky, R. P. Frankenhal, and K. E. Thompson, *J. Electrochem. Soc.* **126**, 2017 (1979).
- ⁶E. Kolawa, F. C. T. So, E. T-S. Pan, and M-A. Nicolet, *Appl. Phys. Lett.* **50**, 854 (1987).
- ⁷L. Krusin-Elbaum, M. Wittmer, and K. S. Yee, *Appl. Phys. Lett.* **50**, 1879 (1987).
- ⁸S. Saito and K. Kuramasu, *Jpn. J. Appl. Phys.* **31**, 135 (1992).
- ⁹J. F. Scott *et al.*, Proceedings of the Eighth IEEE International Symposium on Application of Ferroelectrics, Greenville, S. C., August 1992 (unpublished), p. 356.
- ¹⁰Y. Shimada, *Jpn. J. Appl. Phys.* **35**, 140 (1996).
- ¹¹W. Y. Hsu, *Appl. Phys. Lett.* **66**, 2975 (1995).
- ¹²C. H. Hwang and B. T. Lee, *Korean J. Ceram.* **2**, 95 (1996).
- ¹³W. E. Bell and M. Tagami, *J. Phys. Chem.* **67**, 2432 (1963).
- ¹⁴J. F. Scott, *Integr. Ferroelectr.* **9**, 1 (1995).
- ¹⁵H. B. Michaelson, *J. Appl. Phys.* **48**, 4129 (1977).
- ¹⁶S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981), Chap. 5.
- ¹⁷M. Tomkiewicz, Y. S. Huang, and F. H. Pollak, *J. Electrochem. Soc.* **7**, 1514 (1983).

Equipment simulation of SiGe heteroepitaxy: Model validation by *ab initio* calculations of surface diffusion processes

M. Hierlemann^{a)}

Siemens AG, HL STM, 81730 Munich, Germany

C. Werner and A. Spitzer

Siemens AG, T ME4, 81730 Munich, Germany

(Received 28 May 1997; accepted 30 May 1997)

Incorporation of very little Ge into a Si surface significantly increases deposition during chemical vapor deposition. This is due to the fact that hydrogen and chlorine desorb faster from the SiGe surface making available additional surface sites for adsorption. Two mechanisms are discussed to explain the observed catalytic effect: (i) the diffusion model where surface diffusion of H and Cl atoms from Si to Ge sites opens up an energetically more favorable path for H and Cl desorption via Ge surface sites and (ii) the collective model where incorporation of Ge into Si stimulates an overall change of the electronic structure of the surface, thus leading to increased desorption. *Ab initio* cluster calculations are used in this work to evaluate both models. Binding energies of H and Cl atoms on Si, Ge, and SiGe surfaces are calculated. It is observed that Si-H, Ge-H and Si-Cl, Ge-Cl binding energies do not change whether their neighboring surface atoms are Si atoms or Ge atoms. An overall change of the electronic structure of the surface due to Ge incorporation cannot be observed, making the collective model highly improbable. To evaluate the diffusion model transition states for migration between different surface sites need to be located and the activation barriers need to be calculated. Surface diffusion of H and Cl atoms from Si to Ge is found to be energetically more favorable than desorption of H₂, HCl, or SiCl₂ from Si. Surface diffusion on mixed SiGe surfaces leads to enhanced desorption via Ge surface sites. Thus the diffusion model is considered a valid description. Macroscopic reactor simulations prove that the diffusion model can accurately describe enhanced deposition to explain the observed catalytic effect encountered during growth of SiGe heterolayers. © 1997 American Vacuum Society. [S0734-211X(97)14804-1]

I. INTRODUCTION

Si_{1-x}Ge_x alloys attract considerable attention because of the possibility of band gap engineering with silicon-based materials at much lower cost than with III-V compound semiconductors. Si_{1-x}Ge_x alloys have been proven useful materials for manufacturing high speed electronic devices [High electron mobility transitions (HEMTs), heterojunction bipolar transitions (HBTs)] and optoelectronic components. The composition of the deposited Si_{1-x}Ge_x alloys strongly affects growth during chemical vapor deposition (CVD). With just very little Ge incorporated into the film SiGe deposition is significantly faster than observed during pure Si growth, showing a catalytic behavior of the incorporated Ge. Growth rate and composition of the deposited film are strongly temperature dependent. Better understanding of the SiGe deposition process requires kinetic modeling and simulation of the thermal fluid environment present in a given reactor geometry.

In order to simulate CVD processes, the chemical reaction mechanism and the related reaction rates must be known for the particular deposition chemistry being studied. The chemical mechanism provides a model for the understanding of the underlying chemical rate process. Accurate predictions of observed growth rates indicate that a selected mechanism

is consistent with experimental observations. Chemical mechanisms can represent different levels of accuracy. Simple models consist of only a few lumped rate expressions which have been fitted to experimental data and require modest computational resources. They are not transferable to different reactor geometries. They are not temperature and pressure regimes. Mechanisms based on elementary reactions usually need a higher computational effort. However, they provide a detailed understanding of complex deposition pathways and are valid for a broad range of process conditions.

Formulating mechanisms of elementary gas-phase reactions is feasible since the participating chemical entities are well defined. It is considerably more difficult to estimate surface rate parameters than to estimate gas-phase rate coefficients. The nature of the relevant surface species and their interactions with surface sites and surface atoms are mostly unknown. Computational chemistry has developed rapidly over the last years and allows one to calculate gas-phase molecules as well as molecule-surface interactions. A variety of techniques are available: semiempirical approaches, *ab initio* molecular orbital theory,¹ and density functional methods.² These techniques have been successfully applied to estimate thermodynamic properties and reaction rates for Si-H,³⁻⁶ Si-Cl-H,^{7,8} and Ge-H⁹⁻¹¹ gas-phase systems. Quantum chemistry calculations can also be used to predict surface reaction rates. Cluster calculations as well as computations of extended solids have been demonstrated.¹²⁻¹⁶

^{a)}Electronic mail: matthias.hierlemann@hlistc.siemens.de

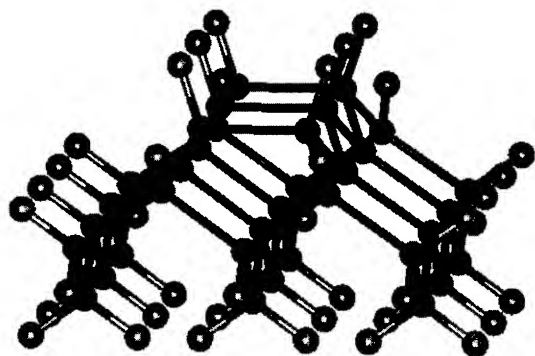


FIG. 1. Si(100)2 \times 1 cluster as used in present calculations. The dangling bonds are saturated with H atoms. The lower two Si layers are fixed during optimization.

In this work cluster calculations are presented to determine binding energies of H and Cl on Si, Ge, and mixed SiGe surfaces. Surface migration of H and Cl adatoms on Si and mixed SiGe surfaces is investigated. The goal of this work is to discriminate between two possible mechanisms for H/Cl-desorption from the mixed SiGe surface during CVD:

1. The diffusion model: It is known that H₂, HCl, and SiCl₂/GeCl₂ desorb faster from Ge than from Si (i.e., activation barriers for H₂-desorption from Ge: 37 kcal/mol,¹⁷⁻¹⁹ from Si: 47 kcal/mol).^{20,21} H and/or Cl atoms diffuse from Si to available Ge sites and desorb from there. The indirect desorption via Ge sites is energetically more favorable than direct desorption from Si sites. The changes in the energetics of desorption are due to local diffusion phenomena.

2. The collective model: Incorporation of Ge into Si stimulates an overall change of the electronic structure of the surface. Incorporation of Ge leads to an increased desorption reactivity (a lowering of the activation barrier) for all Si and Ge sites. The introduced changes in surface energetics are global. They base on collective interactions of a larger number of atoms stimulated by the incorporation of Ge into the surface.

Both mechanisms postulate enhanced desorption from a Ge-containing Si surface, leading to more free surface sites for adsorption of Si and Ge precursors and therefore to higher growth rates. The understanding of this surface catalytic effect is crucial in setting up realistic kinetic models for SiGe alloy deposition.

The following calculations are applied to discriminate between the two suggested mechanisms to explain the catalytic effect:

- (1) Based on a cluster description (the Si cluster is shown in Fig. 1) the binding energies of H and Cl on a Si(100) 2 \times 1 surface, on a Ge(100)2 \times 1 surface, and on a SiGe(100)2 \times 1 surface are calculated.
- (2) The lowest energy barrier (transition state) for surface diffusion of H and Cl atoms from Si to Ge sites is localized.

For the diffusion model to be in effect, the following criteria

have to be fulfilled: (i) the binding energy of H/Cl on Si has to be larger than that of H/Cl on Ge, and (ii) the activation barrier for surface diffusion of H/Cl adatoms from Si to Ge has to be lower than for H₂, HCl or SiCl₂ desorption from Si. The criteria for the collective model are not that straightforward. The possibility of collective effects is evaluated using the following reasoning: if incorporation of Ge stimulates an overall change of the electronic structure of the surface, then the binding energies of Si-H and Si-Cl are expected to change in the presence of a neighboring Ge atom. If this is not the case this means that the binding energies are determined locally and collective effects are highly improbable. The success of this work consists in the elimination of the more unlikely mechanism.

II. METHODOLOGY OF CLUSTER CALCULATIONS

Out of a variety of public domain and commercial codes the MSI BIOSYM software package²² has been chosen to perform the theoretical chemistry calculations. The MSI BIOSYM package combines modules for various calculation techniques [molecular orbital (MO) theory density functional theory (DFT), molecular dynamics (MD)] with comfortable pre- and postprocessing and necessary analysis tools. The following quantum chemistry calculations have been performed using the DFT based DMOL²³ module.

For determining the binding energies of H and Cl adatoms on Si, Ge, and SiGe surfaces and locating the transition states (TS) for lateral H/Cl diffusion, we propose to use a cluster approach. According to Nachtigall and Jordan,¹⁶ cluster calculations lead to reliable activation energies for H-atom diffusion on Si. Nachtigall and Jordan could confirm that cautious cluster calculations are adequate to slab or extended solid calculations. The cluster geometry chosen for our calculations is shown in Fig. 1. The surface is represented by 35 Si and/or Ge atoms. The size of the cluster is selected according to suggestions of Tang and co-workers¹⁴ who investigated the impact of cluster size and cluster relaxation on accuracy. They use four vertical layers of Si atoms to represent the bulk, the 2 \times 1 surface is built by two SiSi dimers [the Si(100) and Ge(100) surfaces are known to reconstruct to a 2 \times 1 structure with surface dimers]. Similar results on a reasonable cluster size are obtained by Nachtigall and Jordan. In our calculations Tangs cluster is extended by one additional surface dimer. Now the middle dimer is neighbored on both sides by other dimers. This way cluster boundary effects (artifacts from describing a continuous surface as a cluster) are prohibited from largely influencing calculations for the center dimer. The systematic error of representing a surface by a finite cluster is minimized. The dangling bonds at the cluster boundaries are saturated with hydrogen. The Ge cluster is built similar to the Si cluster. The mixed SiGe cluster is obtained from the Si cluster by replacing one Si atom of the middle dimer by Ge.

DMOL,²³ an all electron numerical BIOSYM module solving the local density functional equations, is applied to study the Si, Ge, and SiGe surfaces. The DNP basis set which contains 5 atomic basis functions for H, 8 functions for Si and Cl, and

12 functions for Ge is used in our calculations. The DNP basis set uses double-numerical basis functions together with polarization functions, i.e., functions with angular momentum one higher than that of the highest occupied orbital in the free atoms. The DNP basis functions are comparable in quality to Gaussian 6-31G** basis functions and usually yield very reliable calculations. Specific orbitals have been frozen in our calculations: the $1s2s$ for Si and Cl and the $1s2s2p$ for Ge. Closed-shell systems (all bonds are saturated) have been calculated spin-restricted which forces both alpha and beta electrons into the same orbital, whereas open-shell systems (dangling bonds) have been calculated spin-unrestricted to allow electrons with different spins to use different orbitals.

Cluster geometries have been optimized with the lower two layers of Si/Ge atoms frozen at their respective crystal positions. H atoms have been added at the cluster boundaries to saturate the dangling bonds. The bond lengths of the frozen H atoms to the Si/Ge bulk atoms have been taken from calculations on the silane/germane molecule and have been set to 1.48 and 1.52 Å, respectively. Since we are interested in energy differences of certain cluster configurations where the lower two Si layers together with the saturating hydrogens are frozen, the small error introduced by not exactly determining the Si-H/Ge-H bond lengths is canceled. The upper two layers of Si/Ge atoms and the bonded H and Cl atoms were allowed to relax during optimization. Geometry optimization has been done using VWN local spin density functionals. Usage of gradient-corrected functionals for geometry optimization of the given clusters does not further change the optimized geometries and is not necessary. Once the optimum geometry is located improved energies are calculated using the Becke gradient-corrected exchange functional (B₈₈), together with local correlation (VWN) and the gradient-corrected correlation functionals of Perdew and Wang (PW). The gradient-corrected functionals are computationally more expensive than the local functionals, but are necessary to obtain accurate energies. Absolute values for binding energies are expected to be within about 5 kcal/mol of the experimental values for the calculated clusters.

III. CLUSTER CALCULATIONS: RESULTS AND DISCUSSION

A. Binding energies of H and Cl

Binding energies of Si-H/Si-Cl or Ge-H/Ge-Cl for Si, Ge, and SiGe clusters are determined by first optimizing the geometries of the hydrogen saturated clusters and the cluster geometries after abstraction of one H/Cl atom. In a second step the total energies of the optimized structures and a single H/Cl atom are calculated. The binding energy is the difference between the total energy of the saturated cluster and the sum of the total energies of the H/Cl atom and the cluster with one H/Cl abstracted. The calculated binding energies are displayed in Table I together with available values from the literature. Binding energies for Si-H/Si-Cl and

TABLE I. Comparison of calculated binding energies with available literature values.

Structure	Binding energies [kcal/mol]	
	Calculated	Experimental
Molecules		
H-SiH ₃	89.6	90, ^a 92 ^b
H-GeH ₃	84.1	83 ± 2 ^c
Cl-SiCl ₃	100.4	107, ^d 108, ^e 111 ^c
Cl-GeCl ₃	81.7	81 ^f
Si cluster		
Si-H	78.5	82 ^g
Si-Cl	95.0	92 ^h
Ge cluster		
Ge-H	71.0	
Ge-Cl	85.1	
SiGe cluster		
Si-H	78.5	
Ge-H	71.8	
Si-Cl	93.4	
Ge-Cl	82.9	

^aReferences 28 and 29.

^bReference 30.

^cReference 28.

^dReference 29.

^eReferences 31 and 32.

^fReference 33.

^gReferences 34 and 35.

Ge-H/Ge-Cl in SiH₄, SiCl₄, GeH₄, and GeCl₄ molecules have been added as reference to determine the accuracy of the theoretical calculations.

The calculated binding energies of Si-H and Ge-H in the silane and germane molecule exactly reproduce the literature data. The same is true for the binding energy of Ge-Cl in GeCl₄. The accuracy of these calculations is inside the 1 kcal/mol range. The binding energy of Si-Cl in SiCl₄ is lower than all available literature values. The reason for this discrepancy is not clear. The theoretical calculations do well for Si-H and Ge-H systems (see SiH₄, GeH₄), and including Cl does not seem to be a problem either according to our GeCl₄ calculations. Further investigation is needed here. In general we would like to conclude from above molecule calculations that DMOL allows accurate predictions of the energetics of the investigated species.

The cluster calculations that have been performed to determine the binding energies of H/Cl on Si(100)2×1 are able to reproduce in good agreement values available from the literature. The difference between the calculated and the literature values is less than 4 kcal/mol. The larger part of this difference might be attributed to simplifications in the problem setup, e.g., freezing of the lower two Si layers in the cluster. However, this is necessary to keep calculation time inside reasonable limits.

The calculated binding energies of H/Cl on Ge(100)2×1 are lower than of H/Cl on Si(100)2×1. This is expected since the Ge/H bond in germane as well as the Ge-Cl

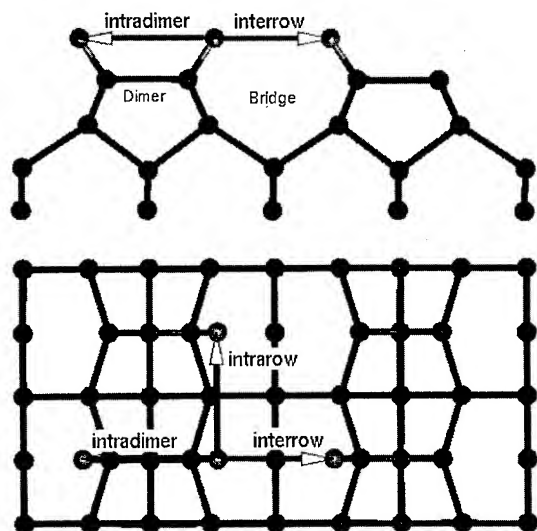


FIG. 2. Three pathways for lateral diffusion are most feasible: (1) migration on the dimer (intradimer), (2) migration between parallel dimers (intrarow), and (3) migration via bridge sites (interrow).

bond in GeCl_4 are weaker than the comparable Si-H and Si-Cl bonds.

The binding energies for Si-H/Si-Cl and Ge-H/Ge-Cl on a mixed $\text{SiGe}(100)2 \times 1$ surface do not differ significantly from those on the pure $\text{Si}(100)2 \times 1$ and $\text{Ge}(100)2 \times 1$ surfaces. The observed energy differences are within the error bars of our calculations. This indicates that the bond strength of Si-H/Si-Cl or Ge-H/Ge-Cl is determined locally. The H and Cl adatoms more or less only "see" its binding partner. From this point of view the collective model seems to be unable to explain enhanced desorption, especially since the observed catalytic effect is most visible at very low Ge content in the SiGe film. Replacement of one Si surface atom by Ge represents very well this low Ge surface concentration. A neighboring Ge atom cannot change significantly the Si-H/Si-Cl binding energies. Hence, a general lowering of the activation barrier for, e.g., H_2 and HCl desorption is highly improbable.

B. Localizing the transition states (TS) for surface hydrogen transfer

1. Surface hydrogen transfer on the $\text{Si}(100)2 \times 1$ surface

Hydrogen surface transfer means that a H adatom migrates from its actual adsorption site to a neighboring one. This requires the actual bond to be broken and the one to the new adsorption site to be formed. It is the activation barrier of such a process we are interested in. On a $\text{Si}(100)2 \times 1$ surface (as well as on a $\text{SiGe}(100)2 \times 1$ surface) every surface H atom has three distinct neighboring adsorption sites as shown in Fig. 2. TSs for the H surface transfer can be found by mapping out the energy surface. This is in general done by placing the H adatom at a number of positions on the surface and minimizing the energy in respect to the distance to the surface. Mapping out all of the energy surface is very

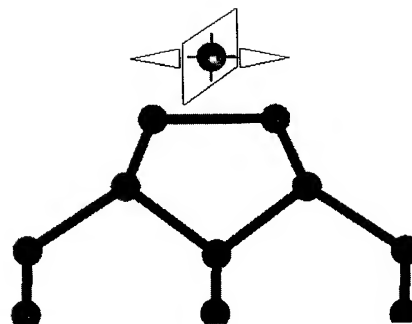


FIG. 3. Energy mapping: the transition state is located by shifting the adatom from one adsorption site to the other. The lateral position between the adsorption states is fixed. However, the adatom can relax in the perpendicular plane.

time consuming. In this work an alternative approach has been used. Three distinct migration pathways have been identified (Fig. 2). The most probable transfer path seems to be the migration of a H atom from one adsorption site of the dimer to the other one (intradimer transfer). This is the shortest possible distance to be travelled by the H adatom. The Si-H/Ge-H bonds do not need to be stretched that much this way. Hence, a low activation energy is expected for H transfer on a dimer. The second shortest path is the transfer between neighboring dimers (intrarow transfer), the longest path is the transfer over a bridge site (interrow transfer).

In a first calculation the TS for migration from one side of the dimer to the other side on a $\text{Si}(100)2 \times 1$ surface has been located. This has been done by stepwise shifting the H adatom from one adsorption site to the other side and calculating the relevant total energy. The H adatom was fixed in its lateral position between the adsorption sites but could relax in the perpendicular plane. This is visualized in Fig. 3. By mapping out the migration path in respect to a maximum energy configuration, the TS (or saddle point on energy surface) could be located. For the Si cluster the saddle point was found right in the middle of the dimer in the same plane as both Si dimer atoms. This is in agreement with calculations of Brocks, Kelly, and Car¹² who located the TS for Si adatom transfer on a dimer exactly in this symmetry position. The calculated activation barrier for the H transfer is 33.5 kcal/mol. This value is lower than the activation energy needed for H_2 desorption from the $\text{Si}(100)2 \times 1$ surface (activation: 47 kcal/mol^{20,21}). Hence, intradimer surface transfer is very probable to occur.

For the second shortest transfer path, migration between two neighboring dimers, the TS has been located in the same way as discussed before. The calculated activation barrier for intrarow H transfer is 46.3 kcal/mol. The activation energy is much higher than for the intradimer transfer. Still, this transfer path is energetically comparable to H_2 desorption. A first calculation of the transfer over a bridge site indicates that interrow transfer is energetically much more difficult. The activation barrier is above 60 kcal/mol. This is significantly higher than the activation energy needed for H_2 desorption. Interrow H transfer is very improbable to occur. The exact TS has not been located.

TABLE II. Calculated transition state structures for diffusion of H and Cl adatoms on Si, Ge, and SiGe surfaces.

	Intradimer H transfer on Si Symmetry plane $E_a = 33.5$ kcal/mole
	Intrarow H transfer on Si Symmetry plane $E_a = 46.3$ kcal/mole
	Intradimer H transfer on SiGe $E_a = 38.7$ kcal/mole (reverse: $E_a = 32.0$ kcal/mole)
	Intradimer Cl transfer on Si Symmetry plane $E_a = 35.4$ kcal/mole
	Intradimer Cl transfer on SiGe Flipping of the dimer in the same direction as observed for H transfer on SiGe is energetically less favourable! $E_a = 37.1$ kcal/mole (reverse: $E_a = 27.3$ kcal/mole)

○ Si ● Ge ● H ● Cl ; distances are in Å

The located TSs are summarized in Table II. Above results are in good agreement with published data (see Ref. 16 and references within).

2. Surface hydrogen transfer on the SiGe(100)2×1 surface

The TS for the mixed SiGe cluster has been located for the energetically most favorable pathway which is the migration from the Si site of the dimer to the Ge site of the dimer or the other way. The calculated TS configuration is displayed in Table II. The energy barrier for H transfer from Si to Ge is calculated to be 38.7 kcal/mol, for H transfer from Ge to Si to be 32.0 kcal/mol. Both values are lower than the activation energies needed for H₂ desorption from Si(100)2×1 (activation: 47 kcal/mol^{20,21}) and Ge(100)2×1 (activation: 37 kcal/mol¹⁷⁻¹⁹). This makes H transfer on SiGe dimers feasible in both directions. Under CVD conditions the equilibrium is shifted in the direction of H transfer from Si to Ge, since the steady state concentration of free Ge sites is observed to be much higher than the concentration of free Si sites. The diffusion model is verified to be a valid approach to explain the observed increased deposition for SiGe heterolayers. Surface diffusion of H adatoms and hydrogen desorp-

tion via Ge is energetically more favourable than direct desorption from Si and results in a higher hydrogen desorption from the SiGe surface.

The activation barrier for H transfer on SiSi dimers is about 5 kcal/mol lower than from Si to Ge on SiGe dimers. This can be explained the following way: for the SiGe dimer the TS is located further away from the Si site than for the SiSi dimer. This is due to the fact that the Ge-H bond is weaker than the Si-H bond. This implies that the Si-H bond has to be stretched further in the TS configuration. As a consequence the activation energy needed to reach the TS configuration is higher.

The calculated values for lateral H diffusion on the SiGe dimer are higher than experimentally estimated ones: according to Russel²⁴ the activation barrier for H transfer from Si to Ge is 23±3 kcal/mol, the barrier for the reverse transfer is approximately 6 kcal/mol. Our theoretical calculations for H transfer on the Si surface agree well with literature data.¹⁶ We have not found problems with the inclusion of Cl atoms into the cluster calculations either. Thus we cannot explain this observed discrepancy. Further investigation is needed.

C. Localizing the transition state (TS) for surface chlorine transfer

1. Surface chlorine transfer on the Si(100)2×1 surface

Intradimer Cl transfer on a pure Si surface has been investigated (for the TS look at Table II). The energy barrier for Cl migration from one site of the dimer to the other one is calculated to be 35.4 kcal/mol. This energy barrier is only a little higher than the one calculated for intradimer H transfer, whereas the binding energy of Cl on Si is significantly higher than that of H on Si. This observed difference can be understood as follows: intradimer Cl transfer requires one Si-Cl bond to be broken and another one to be formed. The Si-Cl bond to be broken is stronger than the equivalent Si-H bond, however the interatomic forces between the unoccupied Si surface atom and the Cl adatom are also higher. The higher forces partly compensate. In addition the Si-Cl bond in the TS configuration does not need to be stretched as much as the Si-H bond since the Si-Cl bond length is larger. Hence we conclude that observed differences in binding energies cannot be applied to estimate the scaling of activation barriers for adatom surface transfer.

2. Surface chlorine transfer on the SiGe(100)2×1 surface

The energy barrier for intradimer Cl transfer from Si to Ge on a mixed SiGe dimer is calculated to be 37.1 kcal/mol. Transfer from Ge to Si requires 27.3 kcal/mol activation. With respect to the 5 kcal/mol accuracy of our calculations these energy barriers are in the same range as for the comparable surface H transfer. Both energy values are significantly below the activation energies needed for desorption of any Cl species (e.g., HCl: $E_{\text{desorp}}=72$ kcal/mol, SiCl₂: $E_{\text{desorp}}=74$ kcal/mol, GeCl₂: $E_{\text{desorp}}=62$ kcal/mol;^{25,26} and references within). The energy differences are much higher

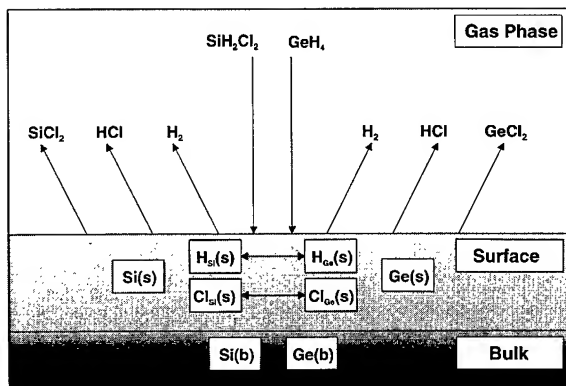


FIG. 4. Schematic representation of the kinetic model for $\text{Si}_{1-x}\text{Ge}_x$ deposition with DCS and GeH_4 .

than between H adatom diffusion and H_2 desorption. This results in an even more increased desorption of Cl species from the SiGe surface relative to the pure Si surface than observed for H_2 desorption due to H surface diffusion. Deposition data confirm this observation. The growth rate enhancement due to incorporation of Ge into the Si surface is more pronounced during SiGe deposition with dichlorosilane and germane than during deposition with silane and germane.

IV. REACTOR SIMULATIONS

From above quantum chemistry calculations we concluded that surface transfer of H and Cl adatoms in between Si and Ge sites can explain faster desorption of H_2 , HCl, and $\text{SiCl}_2/\text{GeCl}_2$.

We want to show next that surface diffusion is the key phenomenon in establishing a working chemical reaction mechanism as it is required for CVD equipment simulation. For this purpose H and Cl surface transfer is implemented into a detailed chemical reaction mechanism for $\text{Si}_{1-x}\text{Ge}_x$ deposition. The chemical reaction mechanism is embedded into calculations for the flow fields and the temperature distribution within a realistic CVD reactor geometry. In the following combined calculations for $\text{Si}_{1-x}\text{Ge}_x$ deposition from dichlorosilane (DCS) and germane (GeH_4) in a single wafer reactor are presented. The calculations are performed using PHOENICS-CVD.²⁷ PHOENICS-CVD accounts for surface species and surface coverages to calculate mass transfer between gas phase, surface, and bulk. Model predictions from our PHOENICS-CVD simulations are compared to in-house experimental data. A schematic representation of the kinetic model used in our calculations is shown in Fig. 4. Previously calculated energies for H and Cl surface transfer have been implemented there. For details on the kinetic model and the experimental reactor set up the reader is referred to Ref. 26.

Figure 5 shows measured and predicted growth rates for various GeH_4 flows. Model predictions agree quite well with the experimental data. Our simulations reproduce very well the strongly enhanced growth for low Ge concentrations. The

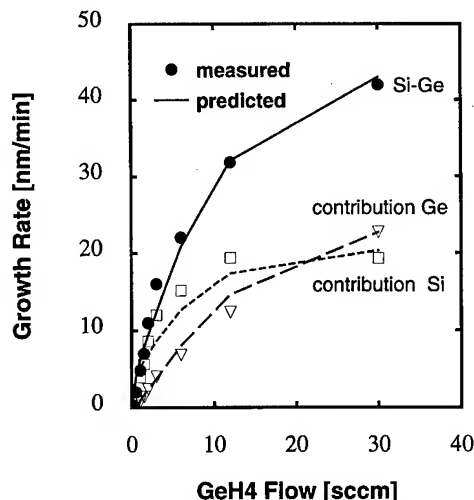


FIG. 5. Measured and predicted growth rates for various GeH_4 flows. Model predictions agree quite well with experimental data.

major increase in deposition at low Ge concentration is mainly due to accelerated Si deposition resulting from enhanced desorption of H_2 , HCl, and $\text{SiCl}_2/\text{GeCl}_2$ via H and Cl surface migration to available Ge sites. Deposition saturates for larger GeH_4 flow rates due to the fact that not the availability of open Ge sites but the reaction rates for surface diffusion limit desorption and hence determine the availability of open Si sites for adsorption. H and Cl surface diffusion is the crucial feature in the reaction mechanism that allows the model to reproduce the observed deposition rates.

V. CONCLUSIONS

This work presents an approach to applying quantum chemistry calculations to surface reaction mechanisms and to explore their contributions to CVD processes. Binding energies of H and Cl on Si, Ge, and mixed SiGe surfaces have been calculated in good agreement with data available from the literature. Si-H, Ge-H and Si-Cl, Ge-Cl binding energies do not change whether their neighboring surface atoms are Si atoms or Ge atoms. An overall change of the electronic structure of the surface due to Ge incorporation cannot be observed. Hence, the collective model cannot explain enhanced deposition. Enhanced deposition is a result of H and Cl surface diffusion from Si to Ge sites and the facilitated H_2 , HCl, and $\text{SiCl}_2/\text{GeCl}_2$ desorption from the mixed SiGe surface. Macroscopic CVD simulations of $\text{Si}_{1-x}\text{Ge}_x$ deposition with DCS and GeH_4 as precursor gases identify H and Cl surface diffusion as the key phenomenon to explain the observed catalytic effect during $\text{Si}_{1-x}\text{Ge}_x$ deposition.

The goal of this work was to discriminate between two possible mechanisms for facilitated H_2 , HCl, and $\text{SiCl}_2/\text{GeCl}_2$ desorption from the mixed SiGe surface. Quantum chemistry calculations provided the necessary information to reach this goal. This result is very encouraging for a further use of quantum chemistry to close the gap between the availability of experimental data and the need for kinetic

information in setting up chemistry models to understand and simulate deposition processes. Quantum chemistry calculations provide access to data where measurements would be much more time consuming, expensive, and sometimes basically not possible. From this point of view there is no alternative to quantum chemistry calculations.

Presented at the 43rd AVS National Symposium, Philadelphia, PA, 14–18 October 1996.

- ¹W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab initio Molecular Orbital Theory* (Wiley, New York, 1986).
- ²R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, The International Series of Monographs on Chemistry, Vol. 16 (Oxford Science, New York, 1989).
- ³C. Sosa and C. Lee, *J. Chem. Phys.* **98**, 8004 (1993).
- ⁴A. F. Sax and J. Kalcher, *J. Phys. Chem.* **95**, 1768 (1991).
- ⁵L. A. Curtiss, K. Raghavachari, P. W. Deutsch, and J. A. Pople, *J. Chem. Phys.* **95**, 2433 (1991).
- ⁶P. Ho, M. E. Coltrin, J. S. Binkley, and C. F. Melius, *J. Phys. Chem.* **90**, 3399 (1986).
- ⁷M. D. Su and H. B. Schlegel, *J. Phys. Chem.* **97**, 9981 (1993).
- ⁸M. D. Su and H. B. Schlegel, *J. Phys. Chem.* **97**, 8732 (1993).
- ⁹R. C. Binning, Jr. and L. A. Curtiss, *J. Chem. Phys.* **92**, 1860 (1990).
- ¹⁰H. Simka, M. Hierlemann, K. F. Jensen, and M. Utz, *J. Electrochem. Soc.* **143**, 2464 (1996), Vol. 5, p. C5-71.
- ¹¹M. Hierlemann, H. Simka, K. F. Jensen, and M. Utz, *Proceedings of EURO-CVD 10, J. Phys. IV*, C5 (1995).
- ¹²G. Brocks, P. J. Kelly, and R. Car, *Phys. Rev. Lett.* **66**, 1729 (1991).
- ¹³G. E. Northrup, *Phys. Rev. B* **44**, 1419 (1991).
- ¹⁴S. Tang, A. J. Freeman, and B. Delley, *Phys. Rev. B* **45**, 1776 (1992).
- ¹⁵A. P. Smith, J. K. Wiggs, H. Jonsson, H. Yan, L. R. Corrales, P. Nachtigall, and K. D. Jordan, *J. Chem. Phys.* **102**, 1044 (1995).
- ¹⁶P. Nachtigall and K. D. Jordan, *J. Chem. Phys.* **102**, 8249 (1995).
- ¹⁷L. Surnev and M. Tikhov, *Surf. Sci.* **138**, 40 (1984).
- ¹⁸R. Malik, E. Gulari, S. H. Li, and P. K. Bhattacharya, *J. Appl. Phys.* **73**, 5193 (1993).
- ¹⁹G. Eres and J. W. Sharp, *J. Vac. Sci. Technol. A* **11**, 2463 (1993).
- ²⁰K. Sinniah, G. M. Sherman, L. B. Lewis, W. H. Weinberg, J. T. Yates, and K. C. Janda, *Phys. Rev. Lett.* **62**, 567 (1989).
- ²¹K. Sinniah, G. M. Sherman, L. B. Lewis, W. H. Weinberg, J. T. Yates, and K. C. Janda, *J. Chem. Phys.* **92**, 5700 (1990).
- ²²Molecular Simulations Incorporated, BIOSYM, Release 3.0.0. San Diego, CA 1995.
- ²³Molecular Simulations Incorporated. DMOL Manual, Release 3.0.0 San Diego, CA, 1995.
- ²⁴N. M. Russel (private communication).
- ²⁵M. Hierlemann, A. Kersch, C. Werner, and H. Schäfer, *J. Electrochem. Soc.* **142**, 295 (1995).
- ²⁶M. Hierlemann and C. Werner, *Proc. Electrochem. Soc.* **96**, 35 (1996).
- ²⁷Heat Concentration and Momentum (CHAM), PHOENICS-CVD, Wimbledon, UK, 1996.
- ²⁸*CRC Handbook of Physics and Chemistry*, edited by D. R. Lide (Chemical Rubber, Boca Raton, 1992), Vol. 73.
- ²⁹R. Walsh, *The Chemistry of Organic Silicon Compounds*, edited by S. Patai and Z. Rappaport (Wiley, New York, 1989), Chap. 5.
- ³⁰P. Gupta, A. C. Dillon, A. S. Bracker, and S. M. George, *Surf. Sci.* **245**, 360 (1991).
- ³¹T. N. Bell, K. A. Perkins, and D. G. Perkins, *J. Chem. Soc. Faraday Trans.* **77**, 1779 (1981).
- ³²R. C. Sausa and A. M. Ronn, *Chem. Phys.* **96**, 183 (1985).
- ³³F. Glocking, *The Chemistry of Germanium* (Academic, New York, 1969), p. 10.
- ³⁴P. Gupta, V. L. Colvin, and S. M. George, *Phys. Rev. B* **37**, 8234 (1988).
- ³⁵M. L. Wise, B. G. Koehler, P. A. Coon, and S. M. George, *Surf. Sci.* **258**, 166 (1991).

Physical and electrical properties in metal-oxide-Si capacitors with various gate electrodes and gate oxides

K.-S. Chang-Liao^{a)} and L.-C. Chen

Department of Nuclear Engineering and Engineering Physics, National Tsing Hua University, Hsinchu, Taiwan, Republic of China

(Received 12 December 1996; accepted 30 May 1997)

The physical and electrical properties in metal-oxide-Si (MOS) capacitors with the gate electrodes deposited using poly-Si or amorphous-Si (*a*-Si) and with the gate oxide grown in O₂ or N₂O have been investigated. The differences of a gate oxide grown in N₂O with a conventional furnace and an oxide film annealed in N₂O by a rapid thermal process (RTP) were also studied. The analysis of physical properties included the thickness variations of oxide films, the shrink ratios of gate electrode films, the nitrogen and hydrogen concentrations in oxides, and the SiO₂/Si interfacial strain. The measurement of electrical properties in MOS capacitors included the interface trap density (*Dit*), the charge-to-breakdown, and the hot electron and radiation induced ΔDit and flatband voltage shifts. To improve the electrical reliability of MOS devices with ultrathin gate oxides, an oxynitride should be introduced although some properties of it are slightly inferior to those of conventional oxides. The combination of a gate electrode deposited using *a*-Si and a gate oxide annealed in N₂O using RTP is shown to exhibit excellent charge-to-breakdown performance and to reduce hot electron and radiation induced interface traps. This improvement can be explained using a mechanism based on the release of compressive stress in the oxide and the relaxation of SiO₂/Si interfacial strain, which could be qualitatively analyzed using infrared spectra. © 1997 American Vacuum Society. [S0734-211X(97)15104-6]

I. INTRODUCTION

Ultrathin (≤ 6 nm) gate dielectrics of high quality are required when metal-oxide-Si (MOS) devices are scaled down for ultralarge-scale integration (ULSI) applications. The long-term reliability of these dielectrics has become an issue as the thickness is decreased leading to a search for a dielectric with better hot-electron hardness than thermal SiO₂.¹ It has been shown that the hardnesses for hot-electron and radiation in MOS devices can be improved by using amorphous Si (*a*-Si) to form the gate electrodes.^{2,3} The improvement is due to the release of compressive stress in the oxide.⁴⁻⁶ Numerous reports have indicated that gate dielectrics grown in N₂O have excellent electrical characteristics such as large charge to breakdown (Q_{BD}), considerable reduction in interface state generation (ΔDit) under constant current stress, and good resistance to radiation exposure.^{1,3,7,8} It has been reported recently that the oxynitride grown in N₂O using rapid thermal process (RTP) shows much better hot-electron hardness than that using a conventional furnace.^{8,9} The mechanism for this improvement is attributed to the relaxation of SiO₂/Si interfacial strain, which is achieved by the nitrogen pileup at the SiO₂/Si interface.⁸⁻¹² The distribution of impurity in the oxide was measured using Fourier transform infrared spectroscopy (FTIR). This technique is frequently used for determining the structure and composition of thin films.^{13,14}

In this work, the physical and electrical properties in MOS capacitors with various gate electrodes and gate oxides were studied. An approach that is a combination of the gate

electrodes deposited using *a*-Si and gate oxides annealed in N₂O using RTP is shown to improve the hot-electron hardness in MOS capacitors. It is found experimentally that the Q_{BD} is markedly increased and the hot-electron and radiation-induced ΔDit are decreased. A possible mechanism for this improvement is based on the release of compressive stress in oxide and the relaxation of SiO₂/Si interfacial strain, which is achieved by the nitrogen pileup at the interface. The SiO₂/Si interfacial strain was qualitatively analyzed by the value of full width at half-maximum (FWHM) at peak position ~ 1080 cm⁻¹ in the infrared spectra. Since the combination of *a*-Si deposited gate-electrode and gate oxide annealed in N₂O by RTP requires few changes to the device fabrication process, it should be useful in practice.

II. EXPERIMENT

The MOS capacitors were fabricated on *p*(100) Si wafers with a diameter of 6 inches and resistivities of 15–25 Ω cm using the conventional poly-Si gate process. Oxide films with a thickness of 6 nm were grown at 900 °C in dry O₂ for 9 min (i.e., “O” samples). Some samples were annealed in N₂O at 1050 °C for 15 sec by RTP (i.e., “OR” samples), which formed an oxide film of 6.4 nm. Oxide film 6.5 nm was grown at 900 °C in N₂O for 65 min (i.e., “N” samples). The oxide thicknesses were measured using an ellipsometer (index=1.46) and were confirmed by the high-frequency capacitance–voltage (*C*–*V*) method. Next, using the conventional low pressure chemical vapor deposition (LPCVD) method, poly-Si films 250 nm thick were deposited at 620 °C on some samples (i.e., “P” samples) and *a*-Si films 330 nm

^{a)}Corresponding author; Electronic mail: lkschang@ne.nthu.edu.tw

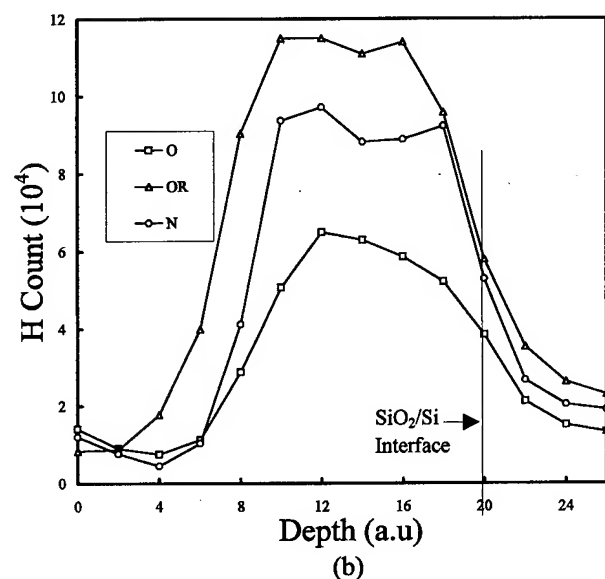
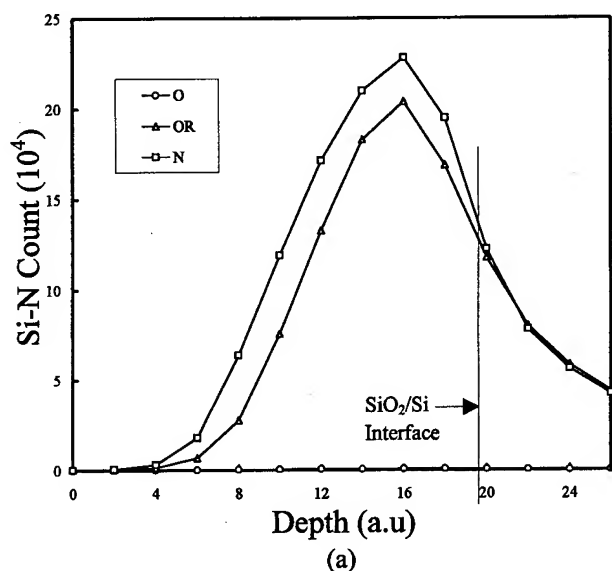


FIG. 1. SIMS results of (a) nitrogen (b) hydrogen in gate oxides for the O, N, and OR samples.

thick were deposited at 550 °C on the remaining samples (i.e., "A" samples). Then, predeposition and drive-in of POCl_3 were performed on all samples at 875 °C for 22 and 20 min, respectively. The sheet resistance of poly-Si was about $50\Omega/\square$ and that of *a*-Si was about $70\Omega/\square$. Finally, after Al films 200 nm thick were evaporated onto the samples and patterned by lithography, the samples were annealed in N_2 at 400 °C for 20 min. The MOS capacitors in this work were devices with various gate electrodes (P or A) and gate oxides (O or N or OR). The gate electrode area of

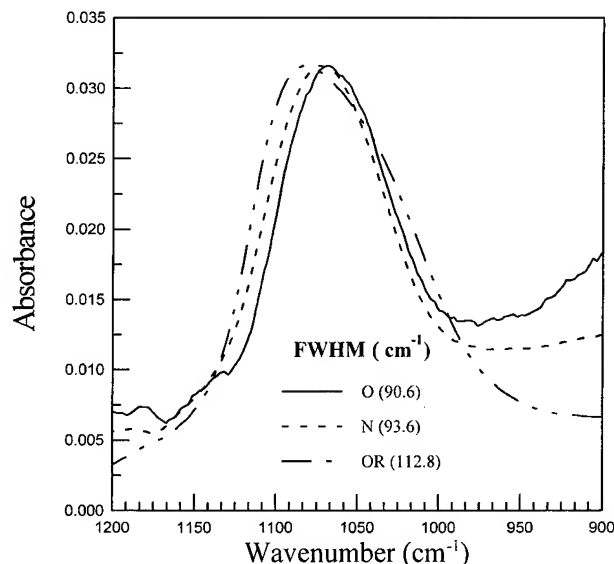


FIG. 2. FTIR spectra of gate oxides for the O, N, and OR samples in the wave number range of 1200–900 cm^{-1} .

the MOS capacitors in this work was $1.257 \times 10^{-3} \text{ cm}^2$. The concentration profiles of nitrogen and hydrogen were analyzed by secondary ion mass spectroscopy (SIMS) using Cs^+ bombardment with an energy of 10 keV. H_2 in the air might interfere with the detection of H atoms. To improve the precision of H measurement by SIMS, the chamber was cooled by liquid nitrogen and the amount of background H_2 minimized. The thicknesses of the poly-Si and *a*-Si films were determined using a surface profile detector, which measures the height and depth of a sample surface. The shrink ratio of the gate electrode film was defined as the value of final thickness (after POCl_3 drive-in)/initial thickness (just after LPCVD). Infrared spectra of the various samples were measured using a FTIR. The time dependent dielectric breakdown (TDDB) was assessed by the Q_{BD} which was measured at a constant current stress of -100 mA/cm^2 . The Q_{BD} values shown in this work were the data from performing over 15 measurements for every sample. The hot-electron stress was performed at a constant current stress of -20 mA/cm^2 . The hot-electron-induced ΔD_{it} and flatband voltage shift (ΔV_{FB}) was determined by the high-frequency and low-frequency (quasistatic) $C-V$ methods.

III. RESULTS AND DISCUSSION

A. Physical properties

The gate oxide thicknesses for O, N, and OR samples are 6.45 ± 0.05 , 6.48 ± 0.16 , and 6.93 ± 0.1 nm, respectively. It can be seen that the data fluctuations of gate oxide thickness

TABLE I. Midgap interface trap density (D_{itm}) for PO, AO, PN, AN, POR, and AOR MOS samples.

Sample	PO	AO	PN	AN	POR	AOR
$D_{\text{itm}} (10^{10} \text{ eV}^{-1} \text{ cm}^{-2})$	6.62 ± 0.37	6.97 ± 0.16	7.11 ± 0.23	7.23 ± 0.29	6.88 ± 0.17	7.04 ± 0.08

TABLE II. Q_{BD} values for the PO, AO, PN, AN, POR, and AOR MOS capacitors.

Sample	PO	AO	PN	AN	POR	AOR
$Q_{BD}(C/cm^2)$	5.98 ± 0.75	6.15 ± 0.92	5.38 ± 0.62	6.56 ± 0.76	7.11 ± 0.43	8.32 ± 0.70

for various samples increase in the order: O, OR, and N. For the MOS devices with ultrathin gate oxide, the variation of oxide thickness should be as small as possible. Therefore, a MOS device with conventional gate oxide (i.e., "O" sample) seems to be good. The gate oxide grown in N_2O by furnace (i.e., "N" sample) is not suitable for the applications of deep submicrometer MOS devices because of its large thickness variation. It has been reported that the thickness of gate oxide grown in N_2O is not so uniform as compared to that grown in O_2 .¹⁵ This may be due to the complexity of decomposition of N_2O at a high temperature. The distribution of oxygen decomposed from N_2O would not be uniform on the surface of wafer.

The gate electrode shrink ratio that is defined as final thickness (after $POCl_3$ drive-in)/initial thickness (just after LPCVD) for PO, AO, PN, AN, POR, and AOR MOS samples are 0.92, 0.76, 0.98, 0.82, 0.96, and 0.78, respectively. The values of shrink ratio of "A" samples (i.e., AO, AN, and AOR) are clearly smaller than those of "P" samples (i.e., PO, PN, and POR). A smaller shrink ratio of gate electrode film would cause a larger increase of the compressive stress in gate oxide, which thereby release the macroscopic strain of this oxide film.¹² It is known that the bending curvature of the SiO_2/Si system is formed because the volume of SiO_2 is twice that of the reacted Si. If a deposited gate electrode film has shrink effects, it would counter the bending of the SiO_2/Si system and then increase the compressive stress in the oxide. It has been shown that the smaller macroscopic strain (intrinsic stress) in the gate oxide film the better hardnesses of hot-electron and radiation in MOS devices.¹⁶ The deposition of α -Si to form gate electrode would be therefore a promising technique in improving the reliability of MOS devices. It is also observed that for the samples with identical gate electrode film and with different gate oxides, the shrink ratio of electrode film decreases in the order: N, OR, and O. The release of macroscopic strain in the gate oxide caused by the gate electrode would be reduced by the incorporation of nitrogen into oxide because the compressive stress at the SiO_2/Si interface is decreased by the nitrogen incorporation.^{6,12}

Figure 1 shows the SIMS results of (a) nitrogen (b) hydrogen in gate oxides for the samples of O, N, and OR. In Fig. 1(a), the count number of nitrogen in the oxide bulk for the N sample is the largest and the centroid of the distribu-

tion is more away from the SiO_2/Si interface. The count number of nitrogen in the OR sample is smaller than in the N sample and the centroid of its distribution appear closer to the SiO_2/Si interface. The count number of nitrogen in the O sample is certainly negligible. It is known that the nitrogen pileup at the SiO_2/Si interface causes some relaxation effects on the interfacial strain.^{8,10} Therefore, the improved hot-electron and radiation hardnesses of MOS capacitors with gate oxide nitrided in N_2O using RTP (i.e., OR sample), which will be shown later in this article, is believed to be due to the relaxation of SiO_2/Si interfacial strain. The nitrogen concentration profile plays an important role in the SiO_2/Si interfacial strain.⁹⁻¹¹ To accomplish the relaxation of SiO_2/Si interfacial strain, the nitrogen concentration in the oxide bulk should be reduced and that at the SiO_2/Si interface should be increased. In Fig. 1(b), the count numbers of hydrogen at the SiO_2/Si interface of various gate oxide samples increase in the order: O, N, and OR. The count numbers of N and OR samples (i.e., oxynitride) are larger than that of O, which may be due to the diffusion of hydrogen-related species into the oxide being enhanced during the process of nitrogen incorporation. It has been reported that the more hydrogen concentration in the oxide, the worse the hot-electron hardness of the MOS device.^{17,18} However, the hydrogen-related species such as Si-H or Si-OH bonds can be minimized by the formation of Si-N bonds.¹⁹ For the oxynitride samples in Fig. 1(b), the large amount of hydrogen in the oxide seems to be not in agreement with the reduction of hydrogen-related bonds which act as electron traps. It is therefore believed that the hydrogen-related electron traps might not be understood only by the concentration of hydrogen atoms in the oxide. This observation needs further investigation.

Figure 2 shows the FTIR spectra of gate oxides for the O, N, and OR samples in the wave number range of 1200–900 cm^{-1} . It is observed that the FWHM for an OR sample has the largest value. The FWHM value of a N sample is smaller than that of OR and it is slightly larger than that of sample O. It has been reported that an increase of the FWHM value indicates an increase of the distorted network in the oxide layer.^{20,21} The higher porosity can be attributed to the increasing disorder in the oxide films. The larger FWHM values for the oxynitride samples (i.e., N and OR samples) could be attributed to the incorporation of nitrogen into the oxide, which might increase the degree of disorder in SiO_2 . Compared with the N sample, the larger FWHM value for OR sample is caused by the increase of the low frequency absorption shoulder. This low frequency shoulder has been shown to be the absorption of Si-O stretching bonds such as Si_2O_3 , SiO , and SiO_2 ,²² whose distributions locate near the SiO_2/Si interface. Therefore, many Si_xO_y species near the

TABLE III. Differences of Q_{BD} for the MOS capacitors with α -Si gate electrodes and those with poly-Si ones, which have the identical gate oxides.

Sample	AO-PO	AN-PN	AOR-POR
$\Delta Q_{BD}(C/cm^2)$	0.17	1.18	1.21

TABLE IV. Q_{BD} values for the PO, AO, PN, AN, POR, and AOR MOS capacitors with gate areas of 1.257×10^{-3} , 3.146×10^{-4} , and $1.767 \times 10^{-4} \text{ cm}^2$, respectively.

$Q_{BD}(\text{C}/\text{cm}^2)$	for sample	PO	AO	PN	AN	POR	AOR
Gate area	1.767	9.88	10.7	10.4	11.3	11.8	12.6
(10^{-4} cm^2)	3.146	9.66	9.78	9.56	10.8	10.8	12.0
	12.57	5.98	6.15	5.38	6.56	7.11	8.32

SiO_2/Si interface could be generated by the nitrogen pileup near the interface as for the OR sample, which then increases the disorder network near the interface. Since the SiO_2/Si interfacial strain is due to the lattice mismatch between SiO_2 and Si,¹² it might be released by increasing the disorder near the interface. On the other hand, only few Si_xO_y species near the SiO_2/Si interface could be produced for the N sample because the nitrogen distribution is somewhat away from the interface, as concluded from the above. Finally, the SiO_2/Si interfacial strain for the N sample can be slightly released as compared with the O sample.

B. Electrical properties

Table I shows the midgap interface trap density (D_{itm}) for PO, AO, PN, AN, POR, and AOR MOS samples. It is observed that the D_{itm} values of the PO samples are the smallest. A MOS device with conventional gate electrode and oxide, which are poly-Si electrode and oxide grown in O_2 by furnace (i.e., PO sample), seem to be fine since a MOS capacitor with a smaller D_{it} indicate the MOS transistor with a larger transconductance. For the samples with identical gate electrode, the D_{itm} values of MOS capacitors with various gate oxides decrease in the order: N, OR, and O. This tendency may be due to the fact that the nitrogen concentration at SiO_2/Si interface decreases in the order: N, OR, and O, which has been shown in Fig. 1(a). It is known that a MOS device with more impurity concentration at SiO_2/Si interface will have a larger D_{it} . For the samples with identical gate oxides, the D_{itm} values of MOS capacitors with gate electrodes of poly-Si are smaller than those of a -Si. This may be because more hydrogen exists at the SiO_2/Si interface for the a -Si sample as compared to the poly-Si due to the large amount of hydrogen in the a -Si film.³

Table II shows the Q_{BD} values for the PO, AO, PN, AN, POR, and AOR MOS capacitors. It is observed that the Q_{BD} values increase in the order: PN, PO, AO, AN, POR, and AOR. Since the Q_{BD} values of the AOR samples are the highest, the time-dependent-dielectric-breakdown (TDDB) performance of MOS capacitors can be significantly improved by a gate electrode deposited using a -Si and a gate

oxide annealed in N_2O using RTP. For the identical gate oxides, it is found that the Q_{BD} values of MOS capacitors with gate electrodes of a -Si are larger than those of poly-Si. This indicates that the TDDB performance of MOS devices can be improved using a -Si to form gate electrodes. The improvement is due to a release of macroscopic strain in the gate oxides caused by an a -Si gate electrode,²⁻⁶ which has been shown above. The Q_{BD} value of the PN sample is lower than that of the PO sample, which has also been observed by other workers.⁸ This may be due to the fact that the large amount of nitrogen in the oxide bulk for the PN sample cannot promote the relaxation of SiO_2/Si interfacial strain and it would finally produce more defects in the oxide. These defects degrade the Q_{BD} performance of PN sample although the hydrogen effects might be suppressed.

For the samples with identical gate oxides, the differences of Q_{BD} for the MOS capacitors with a -Si gate electrodes and those with poly-Si ones are shown in Table III. The positive values of ΔQ_{BD} indicate that the TDDB performance of MOS capacitors with gate electrode of a -Si is better than that of poly-Si. It is worthwhile noting that the ΔQ_{BD} values for gate oxides of oxynitride (i.e., N or OR samples) are much larger than that of conventional oxide (i.e., O sample). This suggests that to significantly improve the TDDB performance of MOS capacitors with a -Si gate electrode, an oxynitride should be used to form the gate oxide. The a -Si gate electrode might increase the hydrogen-related species in the gate oxide,³ and the hydrogen effects can be minimized by the incorporation of nitrogen into the oxide.¹⁹ Therefore, the gate electrode of a -Si should be used together with gate oxide of oxynitride for significantly improving the hot-electron hardness of MOS devices.

Table IV shows the Q_{BD} values for the PO, AO, PN, AN, POR, and AOR MOS capacitors with gate areas of 1.257×10^{-3} , 3.146×10^{-4} , and $1.767 \times 10^{-4} \text{ cm}^2$, respectively. The Q_{BD} values are increased as the gate areas decreased. It has been reported that the hot-electron induced ΔD_{it} of MOS devices are decreased as the gate area decreased.⁴ It is

TABLE V. Midgap interface trap density (ΔD_{itm}) and flatband voltage shift (ΔV_{FB}) for PO, AO, PN, AN, POR, and AOR MOS capacitors which had been subjected to constant current stress with a density of $-20 \text{ mA}/\text{cm}^2$.

Sample	PO	AO	PN	AN	POR	AOR
$\Delta D_{itm}(10^{12} \text{ eV}^{-1} \text{ cm}^{-2})$	8.39	8.11	7.50	6.66	5.34	5.01
$\Delta V_{FB}(\text{V})$	-0.186	-0.302	-0.122	-0.134	-0.254	-0.239

TABLE VI. Midgap interface trap density (ΔD_{it}) and flatband voltage shift (ΔV_{FB}) for PO, AO, PN, AN, POR, and AOR MOS capacitors which had been subjected to radiation exposure of Co-60 gamma ray with a total dose of 2M rad(SiO_2).

Sample	PO	AO	PN	AN	POR	AOR
$\Delta D_{it}(10^{11} \text{ eV}^{-1} \text{ cm}^{-2})$	3.53	2.99	2.46	2.06	2.12	1.76
$\Delta V_{FB}(\text{V})$	-0.023	-0.042	-0.083	-0.075	-0.041	-0.060

known that the Q_{BD} values have a strong dependence on the hot-electron induced ΔD_{it} .²³ The smaller hot-electron induced ΔD_{it} the larger Q_{BD} for the MOS device. Therefore, the gate area effects on the Q_{BD} in Table IV are reasonable. Note that the orders of Q_{BD} performance in various samples have similar tendencies for all gate areas.

Table V shows the midgap interface trap density (ΔD_{it}) and the flatband voltage shift (ΔV_{FB}) for PO, AO, PN, AN, POR, and AOR MOS capacitors which had been subjected to constant current stress with a density of -20 mA/cm^2 for 100 s (2C/cm^2). The ΔD_{it} for the AOR sample has the smallest value. The ΔD_{it} values of MOS capacitors decrease in the order: PO, AO, PN, AN, POR, and AOR. For the samples with identical gate oxides, the ΔD_{it} values of the *a*-Si samples are smaller than those of the poly-Si ones. For the samples with identical gate electrodes and with different gate oxides, the ΔD_{it} values decrease in the order: O, N, and OR. The reduction of hot-electron induced ΔD_{it} may be due to the release of macroscopic strain in the oxide caused by *a*-Si and the relaxation of SiO_2/Si interfacial strain accomplished by the incorporation of nitrogen into the interface. The negative values of ΔV_{FB} indicate that there exist many hole trappings in the oxide. Although lots of electrons are trapped in the oxide under high current density stressing,²⁴ the electrons are easily detrapped and the remaining holes are still trapped in the oxide for gate oxide thickness under $\sim 6 \text{ nm}$.²⁵ The ΔV_{FB} values of MOS samples with gate oxide grown in N_2O by a furnace (i.e., PN and AN) are smaller than those of the other samples. This may be in that the weak Si-H or Si-OH bonds near the SiO_2/Si interface are minimized by the formation of numerous Si-N bonds for the N samples. These strong Si-N bonds can reduce the high-field stress induced charge trapping.¹⁹

Table VI shows the midgap interface trap density (ΔD_{it}) and the flatband voltage shift (ΔV_{FB}) for PO, AO, PN, AN, POR, and AOR MOS capacitors which had been subjected to radiation exposure of Co-60 gamma ray with a total dose of 2M rad(SiO_2). The ΔD_{it} value for the AOR sample has the smallest value. The radiation-induced ΔD_{it} of MOS capacitors can be significantly reduced using *a*-Si gate electrode and gate oxide annealed in N_2O by RTP. The ΔD_{it} values of MOS capacitors decrease in the order: PO, AO, PN, POR, AN, and AOR. These behaviors of various gate electrode/gate oxide samples in Table VI are similar to those in Table V, which thereby have similar mechanism explanations. All the ΔV_{FB} values are negative because the radiation-induced oxide trap charges are positive charges.²⁶ The $|\Delta V_{FB}|$ values of MOS capacitors with the different gate oxides and with the identical gate electrodes increase in the

order: O, OR, and N. The tendency is similar to that of Table I. This could be also due to the defects generations formed by the nitrogen incorporation into the oxide bulk,³ which would increase the radiation-induced oxide-trap-charge (Q_{ot}) in oxynitride samples (i.e., OR and N). For the samples with gate oxides of O or OR, the radiation-induced Q_{ot} in MOS capacitors with gate electrode of *a*-Si are more than those of poly-Si. This might be that the hydrogen effects, enhanced by *a*-Si, would increase the radiation-induced Q_{ot} as explained in Table I.

IV. CONCLUSIONS

Effects of various gate electrodes and gate oxides on the physical and electrical properties in MOS capacitors were discussed. From the thickness variations of gate oxide and the D_{it} of MOS capacitors, the properties of oxynitride samples are slightly worse as compared to the conventional oxide. However, the oxynitride should be used to improve the electrical reliability of MOS devices with ultrathin gate oxides. An oxynitride that was formed by the conventional thermal oxide and then annealed in N_2O by RTP is shown to have superior electrical properties. A gate electrode deposited using *a*-Si would improve the electrical properties of MOS capacitors although it might enhance the hydrogen effects. To significantly improve the hot-electron hardness of MOS capacitors, the gate electrode of *a*-Si should be used together with the gate oxide of oxynitride. The increase of Q_{BD} and the reduction of hot-electron and radiation induced ΔD_{it} can be accomplished by a combination of a gate electrode deposited using *a*-Si and an oxide film annealed in N_2O using RTP. A possible mechanism for this improvement is based on the release of compressive stress in the oxide and the relaxation of SiO_2/Si interfacial strain which is achieved by the nitrogen pileup at the interface. By the infrared spectra analysis of physical properties for the gate oxide, the SiO_2/Si interfacial strain could be qualitatively analyzed by the value of FWHM at peak position $\sim 1080 \text{ cm}^{-1}$. This new approach of gate electrode and gate oxide would be good promise for the ULSI applications.

ACKNOWLEDGMENTS

The authors thank the National Science Council of the Republic of China for supporting this work under Contract No. NSC86-2212-E-007-032. Technical support from the National Nano Devices Laboratory of the ROC is also acknowledged.

- ¹H. S. Momose, T. Morimoto, Y. Ozawa, K. Yamabe, and H. Iwai, IEEE Trans. Electron Devices **41**, 546 (1994).
- ²K. S. Chang-Liao and C. C. Chuang, Electron. Lett. **30**, 1540 (1994).
- ³K. S. Chang-Liao and T. Y. Lan, Jpn. J. Appl. Phys. **1** **35**, L968 (1996).
- ⁴K. S. Chang-Liao and J. G. Hwu, Appl. Phys. Lett. **61**, 675 (1992).
- ⁵K. S. Chang-Liao and J. G. Hwu, IEEE Trans. Electron Devices **41**, 612 (1994).
- ⁶J. T. Yount, P. M. Lenahan, and J. T. Krick, J. Appl. Phys. **76**, 1754 (1994).
- ⁷H. Hwang, W. Ting, B. Maiti, D. L. Kwong, and J. Lee, Appl. Phys. Lett. **57**, 1010 (1990).
- ⁸Y. Okada, P. J. Tobin, V. Lakhota, W. A. Feil, S. A. Ajuria, and R. I. Hegde, Appl. Phys. Lett. **63**, 194 (1993).
- ⁹E. C. Carr, K. A. Ellis, and R. A. Buhrman, Appl. Phys. Lett. **66**, 1492 (1995).
- ¹⁰A. E. Kuiper, H. G. Pomp, and P. M. Asveld, Appl. Phys. Lett. **61**, 1031 (1992).
- ¹¹E. C. Carr and R. A. Buhrman, Appl. Phys. Lett. **63**, 54 (1993).
- ¹²R. P. Vasquez and A. Madhukar, Appl. Phys. Lett. **47**, 998 (1985).
- ¹³I. W. Boyd and J. I. B. Wilson, J. Appl. Phys. **53**, 4166 (1982).
- ¹⁴R. Koba and R. E. Tressler, J. Electrochem. Soc. **135**, 144 (1988).
- ¹⁵T. Y. Chu, W. T. Ting, and D. L. Kwong, J. Electrochem. Soc. **138**, 113 (1991).
- ¹⁶T. B. Hook and T. P. Ma, Appl. Phys. Lett. **48**, 1208 (1986).
- ¹⁷Y. Nissan-Cohen and T. Gorczyca, IEEE Electron Device Lett. **9**, 287 (1988).
- ¹⁸T. Hori and H. Iwasaki, Appl. Phys. Lett. **52**, 736 (1988).
- ¹⁹M. Yasuda, H. Fukuda, T. Iwabuchi, and S. Ohno, Jpn. J. Appl. Phys. **1** **30**, 3597 (1991).
- ²⁰P. Lange, J. Appl. Phys. **66**, 201 (1989).
- ²¹H. J. Schliwinski, U. Schnakenberg, and W. Windbracke, J. Electrochem. Soc. **139**, 1730 (1992).
- ²²K. Sato, J. Electrochem. Soc. **117**, 1065 (1970).
- ²³P. P. Apte and K. C. Saraswat, IEEE Trans. Electron Devices **41**, 1595 (1994).
- ²⁴Z. Liu, H. Wann, P. K. Ko, C. Hu, and Y. C. Cheng, IEEE Electron Device Lett. **13**, 519 (1992).
- ²⁵M. Itsumi, N. Shiono, and M. Shimaya, J. Appl. Phys. **73**, 7515 (1993).
- ²⁶T. P. Ma and P. V. Dressendorfer, *Ionizing Radiation Effects in MOS Devices and Circuits* (Wiley, New York, 1989), Chap. 3.

Gas phase and surface reactions in subatmospheric chemical vapor deposition of tetraethylorthosilicate-ozone

William N. Gill^{a)} and Seshadri Ganguli

Department of Chemical Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180-3590

(Received 18 December 1996; accepted 4 April 1997)

A new physical-chemical model, which applies over a wide range of operating pressures, describes the gas phase and surface reactions in subatmospheric chemical vapor deposition of silicon dioxide for producing inter-layer dielectrics in a cold-wall reactor. Tetraethylorthosilicate (TEOS) reacts in the gas phase to form an intermediate which is adsorbed and reacts on the surface to produce a silicon dioxide film. The results compare favorably with experimental data over a pressure range of 100–600 Torr and a temperature range of 370–500 °C. The concentration distributions of TEOS, intermediate and ozone in the gas phase and their ratios at the surface of the wafer are determined to study gas phase nucleation and the relationship between composition distributions and film quality. Previous models based on low pressure data in the range of 30–90 Torr need to be modified to predict accurately the rates of deposition from 100 Torr to atmospheric pressure. Gas phase reactions cause the maximum in the deposition rates to shift to higher pressures at lower deposition temperatures, both in the model and experiments. At higher pressures, particulates are formed by the gas phase reactions which must be included to represent properly the chemical dynamics of the process. The deposition rate increases up to an asymptotic value as the TEOS flow rate is increased; above this level no further increase in growth rate occurs. This is a consequence of the basic mechanism of the surface reaction which predicts the asymptotic behavior observed. © 1997 American Vacuum Society. [S0734-211X(97)00904-9]

I. INTRODUCTION

New submicron devices with three or four levels of metal and high aspect ratio structures are increasingly difficult to fill without a void. Silicon dioxide deposited as the inter-layer dielectric by the tetraethylorthosilicate-ozone (TEOS-O₃) chemical vapor deposition (CVD) process has the ability to fill high aspect ratio topographies conformally.

Between 30 and 90 Torr TEOS-ozone CVD produces conformal films but they are porous¹ and have a high dielectric constant. At 760 Torr, TEOS-ozone atmospheric pressure chemical vapor deposition (APCVD) provides high quality dielectric films with excellent step coverage.^{2,3} But in APCVD, the deposition rate is too low at temperatures where the film is dense and has a low dielectric constant. Based on these observations, Kim and Gill^{4,5} predicted that SiO₂ can be deposited at relatively high deposition rates, good uniformity and high density film by TEOS-ozone CVD at 200–400 Torr, 400 °C and inlet ozone/TEOS ratio above 5.⁶

Experiments by Shareef *et al.*⁷ conducted in the pressure range between 100 and 600 Torr and the temperature range of 370–500 °C qualitatively confirm the trends predicted by Kim and Gill in deposition rates, uniformity and film density. However, the deposition rates observed in the experiments were much higher than those predicted by the model. This occurred for at least two reasons. First, the reaction rate constants were based on experiments conducted at much lower pressures, between 30 and 90 Torr,⁸ in an Applied Materials Precision 5000 (AM P5000) reactor. Second, it was assumed that the showerhead was maintained at 70 °C. Because the

showerhead of the AM P5000 reactor is not water-cooled, the temperature actually may be as high as 150 °C due to heat transfer by radiation and conduction from the substrate. The activation energies of the gas phase reactions are very high, and the temperature differences are large between wafer and showerhead. Therefore, the reaction rates vary drastically between the showerhead and wafer in this cold-wall reactor. Assuming that the showerhead is at 70 °C leads to rate constants that are too low. Based on the deposition mechanism proposed by Kim and Gill, here we derive new rate constants for the experiments between 100 and 600 Torr in a research reactor with an air-cooled showerhead at 60 °C, which is described in detail elsewhere.^{7,9} The temperature of the wafer surface also was measured accurately using a pyrometer in these new high pressure experiments, and the temperature, pressure, TEOS flow rate and the ozone/TEOS ratio all were varied significantly. These rather large variations in process parameters provide various challenges for the new model to meet as we show here.

II. DEPOSITION MECHANISM

The Kim and Gill model of SiO₂ deposition by TEOS-ozone CVD includes gas phase and surface reactions as indicated in the following description:

(1) Ozone is excited on collision with other molecules and decomposes to atomic and molecular oxygen as proposed by Popovich *et al.*¹⁰ Atomic oxygen recombines with ozone to form molecular oxygen:



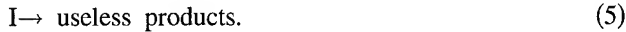
^{a)}Corresponding author; Electronic mail: gillw@rpi.edu



(2) TEOS is oxidized by excited ozone to an intermediate specie (I) which is assumed to be silanol. Fourier transform infrared (FTIR) studies of TEOS-O₃ CVD by Kawahara *et al.*¹¹ also suggest that this occurs as



(3) Some fraction of the intermediate specie decomposes to useless products in the gas phase before reaching the wafer surface



(4) Undecomposed intermediate species adsorbs on the surface, and reacts on the surface according to the Langmuir-Hinshelwood mechanism to produce silicon dioxide. It is assumed that the decomposition reaction is the rate limiting step



A similar mechanism was proposed by Desu¹² for SiO₂ deposition in the absence of ozone from the thermal decomposition of TEOS at high temperatures.

If we assume steady state for the atomic oxygen and excited ozone concentrations and dilute concentrations of reactants, then the rates of the gas phase reactions of TEOS-O₃ CVD are:⁵

$$R_{g,\text{O}_3} = -k_{\text{O}_3} C_{\text{M}} C_{\text{O}_3}, \quad (8)$$

$$R_{g,\text{TEOS}} = -k_{\text{TEOS}} C_{\text{M}} C_{\text{O}_3} C_{\text{TEOS}}, \quad (9)$$

$$R_{g,\text{I}} = k_{\text{TEOS}} C_{\text{M}} C_{\text{O}_3} C_{\text{TEOS}} - k_{\text{I}} C_{\text{I}}. \quad (10)$$

Figure 5 of Kim and Gill⁵ indicates clearly the existence of gas phase reactions and the consequence of their role is discussed on the basis of experimental behavior by Shareef *et al.*⁷

The rate of deposition of SiO₂ from the surface reaction mechanism is derived by the Hougen-Watson-Langmuir-Hinshelwood method to be

$$R_{\text{SiO}_2} = \frac{k_r K_a C_{\text{I}}(0)}{1 + K_a C_{\text{I}}(0)}, \quad (11)$$

where K_a is the equilibrium constant for the adsorption of intermediate, k_r is the surface reaction rate and $C_{\text{I}}(0)$ is the concentration of the intermediate at the gas-wafer interface. At high concentration of intermediate $C_{\text{I}}(0)$, Equation (11) suggests that the rate becomes constant and equal to k_r . We assume in obtaining Equation (11) that the concentration of the by-product formed on an un-patterned wafer by the surface reaction is small and does not affect the deposition rate. This assumption is not true in high aspect ratio microstructures.¹³ Note that the ozone decomposition in the gas phase is given by Eq. (8) and is not influenced by the

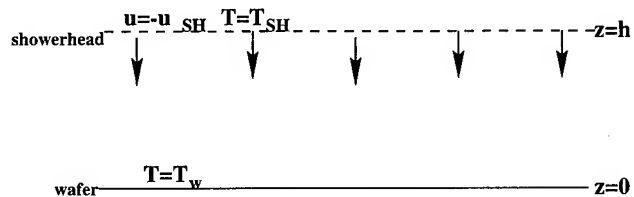


FIG. 1. Schematic of the CVD reactor used in the analysis.

TEOS concentration. The same behavior for ozone decomposition was observed by Kawahara *et al.*¹¹ in their experiments.

III. MODEL DESCRIPTION

To determine the kinetic coefficients it is desirable to simplify the fluid mechanics and thermal models to make the problem tractable. Reference 9 describes the stagnation point flow research reactor in which the experiments we analyze here were carried out. In such a system, a large part of the wafer surface is uniformly accessible and can be analyzed by a one dimensional model¹⁴ as shown in Figure 1. The vertical velocity in the gap between showerhead and wafer is approximated sufficiently well by a linear profile, $u(z) = -u_{\text{SH}}(z/h)$ where z is the distance from the wafer and h is the distance between the showerhead and the wafer. The heat transfer in the gas phase occurs mainly by conduction and the temperature can also be approximated as

$$\frac{T(z) - T_w}{T_{\text{SH}} - T_w} = \frac{z}{h}. \quad (12)$$

The concentrations of the different species ($i = \text{O}_3$, TEOS, and I) are determined by the convective-diffusion equation:

$$\frac{d}{dz} \left(D_i \frac{dC_i}{dz} \right) + \frac{d}{dz} \left(D_i C_i \alpha_i \frac{d \ln T}{dz} \right) + R_{g,i} = u \frac{dC_i}{dz}, \quad (13)$$

where D_i is the diffusion coefficient, α_i is the thermal diffusion factor of specie i , and the coordinate system is shown in Figure 1.

The boundary conditions for the different species ($i = \text{O}_3$ and TEOS) are

$$D_i \frac{dC_i}{dz}(0) + D_i \alpha_i C_i \frac{d \ln T}{dz}(0) = 0 \quad (14)$$

at the wafer, and

$$u_{\text{SH}} C_i(h) + D_i \frac{dC_i}{dz}(h) = u_{\text{SH}} C_{i,\text{in}} \quad (15)$$

at the showerhead where $C_{i,\text{in}}$ is the known inlet concentration of the i th species. For the intermediate specie I, the boundary conditions at a wafer and showerhead are

$$D_{\text{I}} \frac{dC_{\text{I}}}{dz}(0) + D_{\text{I}} \alpha_{\text{I}} C_{\text{I}} \frac{d \ln T}{dz}(0) = \frac{k_r K_a C_{\text{I}}(0)}{1 + K_a C_{\text{I}}(0)}, \quad (16)$$

$$u_{\text{SH}} C_{\text{I}}(h) + D_{\text{I}} \frac{dC_{\text{I}}}{dz}(h) = 0. \quad (17)$$

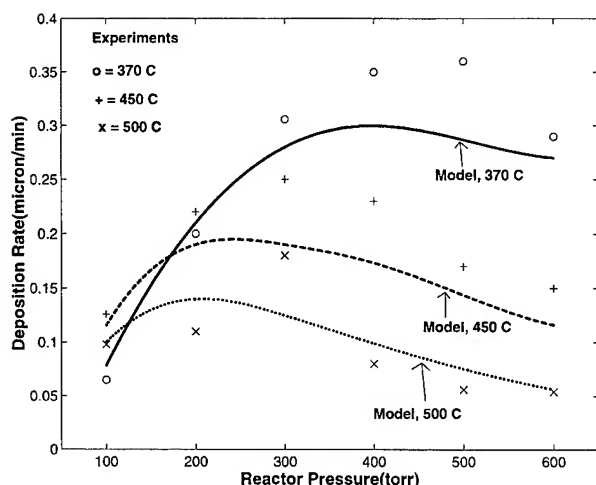


FIG. 2. Comparison of model with experimental data. Flow rates: $N_2=20\,000$ sccm, $O_2=10\,000$ sccm, $O_3=325$ sccm, TEOS=41 sccm. Experimental data for the figure are obtained from Ref. 17.

At the showerhead, we use the Danckwerts condition¹⁵ which applies when analyzing CVD processes where diffusion is important in the transport of the species.⁵ Equation (16) states that the rate of diffusion of intermediate at the surface is equal to the rate of reaction. This set of differential equations in dimensionless form was solved numerically using the D02RAF subroutine of the Numerical Algorithm Group (NAG) library¹⁶ to determine systematically the kinetic coefficients that are consistent globally with the various results obtained experimentally. The Appendix gives the procedure to determine the physical properties, α_i and D_i .

IV. RESULTS AND DISCUSSION

A. Comparison between model predictions and experiments

A systematic procedure based on the mechanism of the gas phase and surface reactions was used to extract the rate constants from the experimental data. Evaluation of the surface reaction rate k_r was based on the experimental deposition rate data attaining an asymptotic value at high TEOS flow rates because the film surface is saturated by the intermediate specie and the resistance to adsorption of the intermediate is negligible. Thus the growth rate is limited only by the surface reaction rate. Consequently the term $K_a C_I(0)$ in the denominator of Equation (11) is much larger than 1, and the asymptotic value of the growth rate as shown in Figure 5 can be approximated as $R_{SiO_2} \approx k_r$. With k_r known, an iterative procedure was used to determine the pre-exponential factor and activation energy of the other three rate constants and K_a that characterize the TEOS- O_3 deposition process. Initial values for the different rate constants were obtained from Kim and Gill⁵ based on data from 30 to 90 Torr. The final values obtained for the rate constants over the range of 100–600 Torr, 370–500 °C, $N_2=20\,000$ sccm, $O_2=10\,000$ sccm, $O_3=325$ sccm, TEOS=41 sccm were:

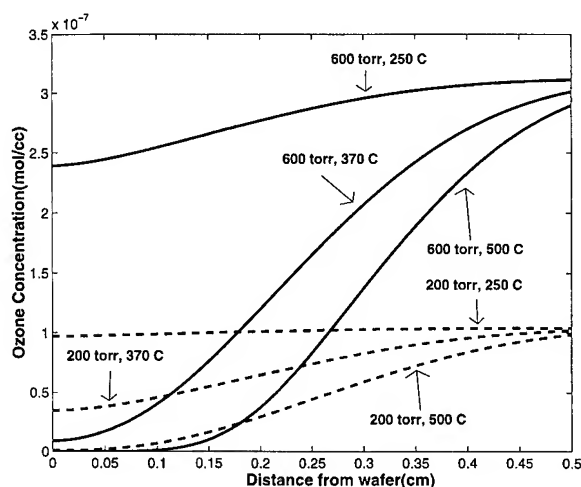


FIG. 3. Concentration distribution of ozone between the showerhead and the wafer surface. Flow rates: $N_2=20\,000$ sccm, $O_2=10\,000$ sccm, $O_3=325$ sccm, TEOS=41 sccm. The solid and dashed lines are for 600 and 200 Torr, respectively.

$$k_{O_3} = 1.01 \times 10^{16} \exp(-24.9/RT) \text{ cm}^3 \text{ mol}^{-1} \text{ s}^{-1},$$

$$k_{TEOS} = 1.06 \times 10^{21} \exp(-19.6/RT) \text{ cm}^6 \text{ mol}^{-2} \text{ s}^{-1},$$

$$k_I = 4.12 \times 10^5 \exp(-11.0/RT) \text{ s}^{-1},$$

$$k_r = 6.38 \times 10^{-7} \exp(-4.31/RT) \text{ cm}^{-2} \text{ mol s}^{-1},$$

and

$$K_a = 5.02 \times 10^{11} \exp(-4.1/RT) \text{ cm}^3 \text{ mol}^{-1},$$

where R is universal gas constant = 0.002 kcal/mol K. These values were obtained by requiring the model to satisfy simultaneously various criteria with respect to behavior of the experimental data that will be discussed next.

The growth rate depends on the concentration of intermediate and the temperature at the wafer surface. However the wafer temperature also affects the temperature at various points in the gas phase and therefore the gas phase reactions. Consequently, the temperature and concentration of intermediate at the wafer surface are intimately connected through the gas phase reactions. A comparison of the predictions of the model and the experimental data is shown in Figure 2. The model predicts correctly the increase-decrease behavior of the deposition rate as the total reactor pressure or the wafer temperature is increased. The temperature effect is seen in the crossover of the curves between 100 and 200 Torr. Also, for all three wafer temperatures the rate increases to a maximum and then decreases, and the maximum occurs at higher pressure with lower temperature. This behavior is explained in Figures 3 and 4 and Table I. The increase in deposition rate from row 1 to 2 of Table I occurs because the gas phase temperature is higher at higher wafer temperatures and more TEOS is oxidized to form the intermediate. However, even though more intermediate is produced in the gas phase at 500 °C and 200 Torr than at 370 °C and 200 Torr, the decomposition reaction given by Equation (5) consumes it rapidly at higher temperatures before it reaches the wafer.

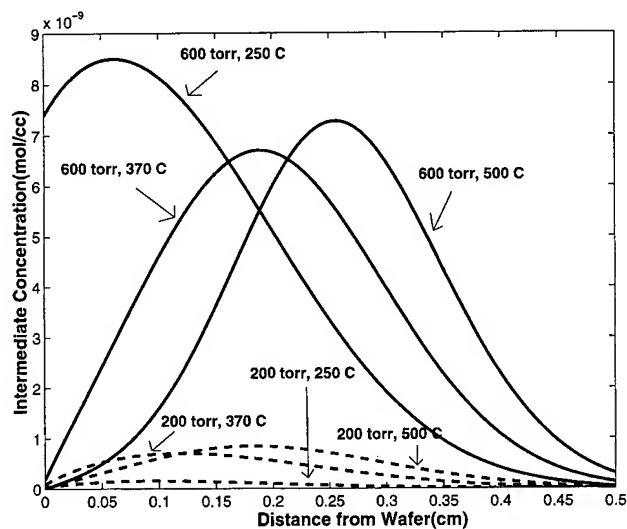


FIG. 4. Concentration distribution of intermediate between the showerhead and the wafer surface. Flow rates: $N_2=20\,000$ sccm, $O_2=10\,000$ sccm, $O_3=325$ sccm, TEOS=41 sccm. The solid and dashed lines are for 600 and 200 Torr, respectively.

The depletion of intermediate is enhanced further by rapid consumption of ozone at these high temperatures (Figure 3). Therefore the concentrations on the wafer surface and the rate of film growth are lower at 500 °C.

As the reactor pressure is increased, the concentrations of the reactants in the gas phase become higher and more intermediate is formed in the reactor. In Table I, we see that $C_{I_{max}}$ at 600 Torr is an order of magnitude higher than $C_{I_{max}}$ at 200 Torr. Hence the deposition rate in rows 4 and 5 is higher than in rows 1 and 2 of Table I. The deposition rate at 600 Torr and 500 °C is very low because of the increased rate of ozone consumption so that very little TEOS is oxidized to form intermediate close to the wafer surface.

As explained above, the combined effect of temperature and pressure leads to a shift in the maximum deposition rate to lower pressures at high temperature. The maximum amount of intermediate in Table I is formed at 250 °C and 600 Torr. At this low wafer temperature the intermediate concentration maximum occurs close to the wafer (no significant ozone consumption or intermediate decomposition)

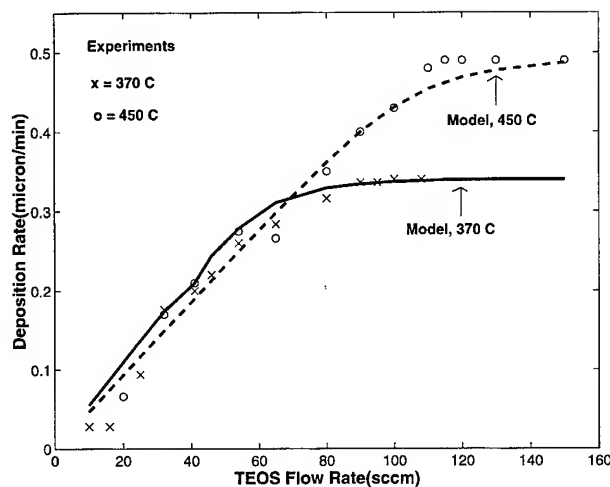


FIG. 5. Deposition rate versus TEOS flow rate showing that the deposition rate reaches an asymptotic value at high flow rates. Pressure = 200 Torr, flow rates: $N_2=20\,000$ sccm, $O_2=10\,000$ sccm, $O_3=325$ sccm. Experimental data for the figure are obtained from Ref. 17.

and the high pressure provides high reactant concentrations. However, the deposition rate is limited by Equation (11) to k_r and the maximum deposition rate possible at a 250 °C wafer temperature is ~ 0.16 $\mu\text{m}/\text{min}$ because of this saturation effect.

In experiments at high reactor pressures and temperatures, white particulate residues were observed on the reactor windows due to the intermediate decomposition reactions in the gas phase. With an increase in pressure, more intermediate is formed and is decomposed at high wafer temperatures as shown in Figure 4; this leads to particulate formation. The observation of particulate formation further supports the parasitic gas phase reaction postulated by Kim and Gill.^{4,5}

Figure 5 shows a comparison of the model predictions with the experimental data for deposition rate at various TEOS flow rates. The predictions agree well with the experiments. The film growth rates reach asymptotic values at high TEOS flow rates, because the wafer surface becomes saturated with intermediate. Under these conditions, the growth rate is determined only by the surface reaction rate constant, k_r .

TABLE I. The variation of intermediate concentration with the gap spacing for different process conditions and its effect on deposition rate.

No.	Process conditions		Values where $C_{I_{max}}$ occur			C_I at wafer surface ($\text{mol}/\text{cc} \times 10^{10}$)	Growth rate ($\mu\text{m}/\text{min}$)
	Reactor pressure (Torr)	Wafer temperature (°C)	Distance from wafer (cm)	Gas phase temperature (°C)	$C_{I_{max}}$ ($\text{mol}/\text{cc} \times 10^{10}$)		
1	200	250	0.1	213	1.4	0.30	0.04
2		370	0.12	299	6.7	0.83	0.22
3		500	0.18	342	8.0	0.08	0.14
4	600	250	0.06	227	78	66	0.16
5		370	0.19	255	63	1.6	0.27
6		500	0.26	276	69	0.03	0.06

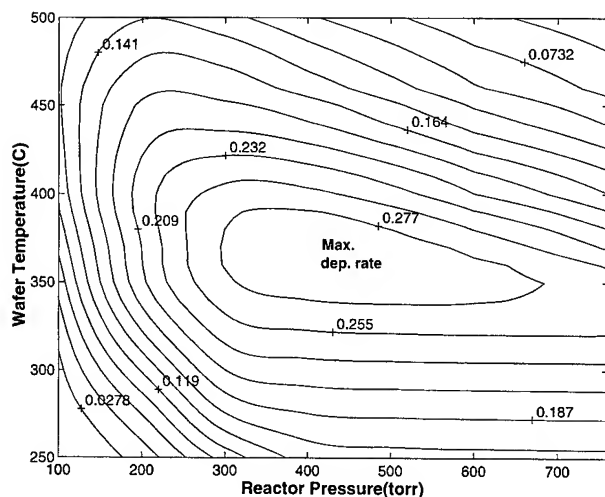


FIG. 6. A contour plot which shows that the maximum deposition rate of $0.29 \mu\text{m/min}$ occurs at reactor pressure = 450 Torr and wafer temperature = 375°C for a TEOS and ozone flow rates of 40 sccm and 325 sccm, respectively. $\text{N}_2 = 20\,000$ sccm and $\text{O}_2 = 10\,000$ sccm. Numbers in the figure represent deposition rate in $\mu\text{m/min}$.

B. Effect of process variables on deposition rate

In a single-wafer reactor as described in Ref. 9, the process variables that can be manipulated to change the deposition process are: (1) reactor pressure, (2) wafer temperature, (3) showerhead temperature, (4) carrier gas flow rate, (5) ozone flow rate, (6) TEOS flow rate, and (7) gap spacing between the wafer and showerhead. The model developed in this paper can be used to determine the effect of these process variables on the deposition rate. The question we answer here is, what combination of temperature and pressure gives a maximum rate of deposition for various fixed values of the flow rates of reactants.

Extensive simulations were conducted to obtain the reactor pressure and wafer temperature for which the deposition rate is a maximum at various flow rates of TEOS and ozone. Figure 6 shows a contour plot of the deposition rate over ranges of values of pressure (100–760 Torr) and wafer temperature (250–500 $^\circ\text{C}$) for TEOS and ozone flow rates of 40 and 325 sccm, respectively. We see that the maximum deposition rate of $0.29 \mu\text{m/min}$ occurs at a reactor pressure of 450 Torr and a wafer temperature of 375°C . The results of similar calculations give the temperatures and pressures at which the deposition rates are a maximum for other flow rates of TEOS and ozone and these are summarized in Table II. For lower flow rates of TEOS or ozone, R_{max} , the deposition rate maximum occurs at a higher reactor pressure and lower wafer temperature. These changes in process conditions lead to higher concentrations of TEOS and ozone in the reactor which compensates for the decreased concentration of reactants due to the lowering of their flow rates. The temperature and pressure for R_{max} changes with different flow rates of TEOS and O_3 , and one needs to find the combination of processing condition which leads to R_{max} . The maximum deposition rate increases as the TEOS flow rate is increased. But the slope of the deposition rate decreases as the TEOS

TABLE II. Regions of maximum deposition rate for various flow rates of TEOS and ozone. Flow rates: $\text{N}_2 = 20\,000$ sccm and $\text{O}_3 = 10\,000$ sccm. Gap spacing = 0.5 cm.

No.	Reactant flow rates			Region of maximum deposition rate		
	TEOS (sccm)	Ozone (sccm)	O_3/TEOS	Pressure (Torr)	Temperature ($^\circ\text{C}$)	Max. dep. rate ($\mu\text{m/min}$)
1	20	325	16	600	300	0.18
2	40	325	8	450	375	0.29
3	60	325	5.5	375	400	0.35
4	80	325	4	300	425	0.39
5	120	325	2.7	200	450	0.46
6	40	650	16	300	390	0.32
7	80	650	8	225	450	0.46

flow rate increases because the wafer surface becomes saturated with the intermediate (Figure 5). This limits the cost-effectiveness of increasing the TEOS flow rate in the reactor. Moreover, if the O_3 to TEOS ratio is too low, the quality of the film is poor.^{2,6} The increase in deposition rate with increasing ozone flow rate is not very high as shown in Table II. But a high flow rate should be maintained to obtain good film quality. The ozone flow rate is limited by the total oxygen flow rate and the efficiency of the ozonator which is usually less than 8% by volume.

Shareef¹⁷ characterized the quality of the silicon dioxide film deposited at 400°C and 100–400 Torr by TEOS-ozone CVD. The measurements of dielectric constant and density of the film show that at reactor pressures above 300 Torr, the quality of the film is comparable to thermally grown oxide. Simulations were performed to obtain the TEOS and ozone flow rate for which a deposition rate maximum occurs at 400°C and 400 Torr. Figure 7 shows the deposition rate reaches an asymptotic value of $0.4 \mu\text{m/min}$ at 400°C and 400 Torr for TEOS and ozone flow rates of 80 sccm and 600 sccm, respectively. The simulation results together with experiments of Shareef¹⁷ show that the TEOS- O_3 CVD system

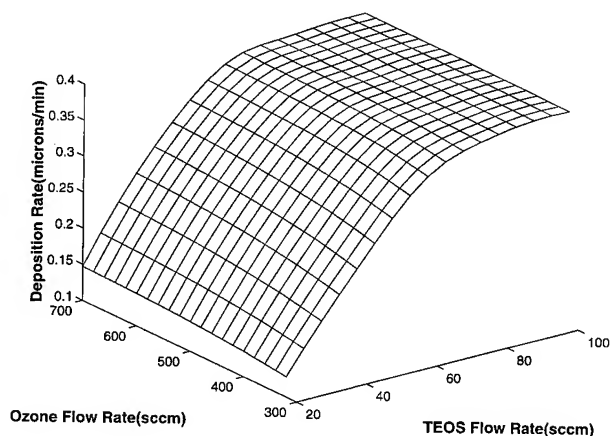


FIG. 7. Deposition rate predicted by model at 400°C and 400 Torr for various flow rates of TEOS and ozone. Flow rates: $\text{N}_2 = 20\,000$ sccm, $\text{O}_2 = 10\,000$ sccm.

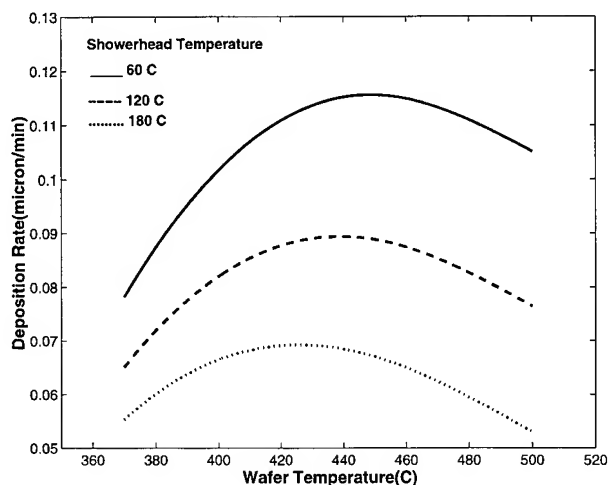


FIG. 8. A plot showing model prediction of the decrease in deposition rate as the showerhead temperature is raised. Pressure = 100 Torr, Flow rates: $N_2=20\,000$ sccm, $O_2=10\,000$ sccm, $O_3=325$ sccm.

can be used to deposit inter-layer dielectric at growth rates of $0.4\ \mu\text{m}/\text{min}$ with film quality comparable to thermally grown oxide.

Figure 8 shows that higher showerhead temperatures decrease the deposition rate drastically because more intermediate and ozone decomposition occurs closer to the showerhead. Thus particulate formation also may occur in the gas phase if the temperature of the showerhead is too high. Thus the temperature of the showerhead of a TEOS- O_3 CVD reactor should be low, so that maximum intermediate formation occurs close to the wafer surface.

The showerhead of CVD reactors usually has slots or small holes, and the total open area of the showerhead is much smaller than the reactor area. The gas mixture thus enters the reactor as jets, but these jets should mix within a small fraction of the gap spacing to obtain a uniform film on the wafer surface. Moreover, re-circulation by natural convection may occur in these reactors at large gap spacing and high pressure and this increases the nonuniformity of the film thickness. Therefore, to study the effect of carrier gas flow rate and gap spacing, a more complex model which simulates the jet phenomena and natural convection is required. For a total flow rate of $30\,000$ sccm and gap spacing of 0.5 cm, experiments showed that the film deposited is uniform and the jet effect or re-circulation does not affect film uniformity. Consequently a linear velocity distribution between the showerhead and wafer is a good assumption in this case.

V. CONCLUSIONS

Based on the model results, we can conclude that:

- (1) The kinetic model proposed in this article successfully describes the deposition characteristics of TEOS- O_3 CVD in the ranges of 100–600 Torr and 370–500 °C.
- (2) The decrease of the deposition rate at higher reactor temperatures or pressures occurs due to depletion of the ozone concentration near the wafer surface and increased parasitic gas phase reaction of the intermediate.

- (3) At high TEOS flow rates, the deposition rate reaches an asymptotic value because the wafer surface becomes saturated with the intermediate specie.
- (4) High quality films can be produced at high deposition rates in the neighborhood of 400 °C and 400 Torr.
- (5) The showerhead should be cooled to improve the deposition rate as shown in Figure 8.

APPENDIX

The physical properties of the gas mixtures were determined at a pressure of 1.0 Torr and 343 K as follows:

Reactant gases	Parameters	Carrier gases		
		O_2	N_2	He
O_3	D^o	139.1	134.7	513.2
	A	0.317	0.313	1.24
	B	55.72	50.8	67.32
TEOS	D^o	57.6	57.07	248.3
	A	1.19	0.912	2.61
	B	196.8	333.8	224.4
I	D^o	82.03	80.48	347.1
	A	0.808	0.654	1.87
	B	203.6	152.5	146

$$\frac{1}{D_k} = \frac{1}{\sum_j x_j D_{kj}}, \quad (\text{A1})$$

where $k = O_3$, TEOS or I and $j = O_2$, N_2 or He. D_{kj} is the binary diffusion coefficient in cm^2/s of the k th species in j and x_j represents the mole fraction of the j th species. The diffusion coefficient at any other temperature T (in K), and pressure P (in Torr) is given by

$$D_{kj}(T, P) = D_{kj}^o \frac{T^{1.5}}{6352.4P}. \quad (\text{A2})$$

The thermal diffusion coefficient was calculated as

$$\alpha_k = A_k - \frac{B_k}{T}, \quad (\text{A3})$$

where

$$A_k = \sum_j x_j A_{kj} \quad (\text{A4})$$

and

$$B_k = \sum_j x_j B_{kj}, \quad (\text{A5})$$

where A_{kj} is dimensionless and B_{kj} has the units of temperature.

ACKNOWLEDGMENTS

The work was supported financially in part by New York Sematech Center of Excellence at Rensselaer Polytechnic Institute, SRC Contact No. 95-MC-508, and by New York State Energy Research and Development Authority. The val-

ues of the physical properties of different species and the gas mixture were evaluated using correlations derived by Dr. E. J. Kim. The authors would also like to thank Satish Lakshmanan for reviewing this manuscript.

¹S. Nguyen, D. Dobuzinsky, D. Harmon, R. Gleason, and S. Fridman, *J. Electrochem. Soc.* **13**, 2209 (1990).

²H. Kotani, M. Matsuura, A. Fujji, H. Genjou, and S. Naguo, *IEEE IEDM* **89**, 669 (1989).

³Y. Ikeda, Y. Numasawa, and M. Sakamoto, *J. Electron. Mater.* **19**, 45 (1990).

⁴E. J. Kim and W. N. Gill, *J. Cryst. Growth* **140**, 308 (1994).

⁵E. J. Kim and W. N. Gill, *J. Cryst. Growth* **140**, 315 (1994).

⁶E. J. Kim and W. N. Gill, *J. Electrochem. Soc.* **141**, 3462 (1994).

⁷I. A. Shareef, G. W. Rubloff, and W. N. Gill, *J. Vac. Sci. Technol. B* **14**, 772 (1996).

⁸S. Fugardi, MS thesis, Rensselaer Polytechnic Institute, Troy, NY, 1989.

⁹I. A. Shareef, G. W. Rubloff, M. Anderle, W. N. Gill, J. Cotte, and D. H. Kim, *J. Vac. Sci. Technol. B* **13**, 1888 (1995).

¹⁰M. O. Popovich, G. V. Egorova, and Yu. V. Filippov, *Russ. J. Phys. Chem.* **59**, 165 (1985).

¹¹T. Kawahara, A. Yuuki, and Y. Matsui, *Jpn. J. Appl. Phys.* **1** **31**, 2925 (1992).

¹²S. B. Desu, *J. Am. Ceram. Soc.* **72**, 1615 (1989).

¹³R. J. Soave, S. Ganguli, W. N. Gill, Y. Shacham-Diamand, and J. W. Mayer, *Appl. Phys. Lett.* **67**, 3286 (1995).

¹⁴C. Houtman, D. B. Graves, and K. F. Jensen, *J. Electrochem. Soc.* **133**, 961 (1986).

¹⁵P. V. Danckwerts, *Chem. Eng. Sci.* **2**, 1(1953).

¹⁶Software of Numerical Algorithms Group Inc., Downers Grove, IL.

¹⁷I. A. Shareef, PhD thesis, Rensselaer Polytechnic Institute, Troy, NY, 1995.

Testing of a rapid fault detection model for quality control: Borophosphosilicate glass thin films monitored by infrared absorption spectroscopy

S. Zhang, J. E. Franke, and T. M. Niemczyk^{a)}

Department of Chemistry, University of New Mexico, Albuquerque, New Mexico 87131

D. M. Haaland

Sandia National Laboratories, Albuquerque, New Mexico 87185-0342

J. N. Cox and I. Banerjee

Intel Corporation, 3065 Bowers Avenue, Santa Clara, California 95052

(Received 3 January 1997; accepted 4 April 1997)

Infrared absorption spectra of 108 borophosphosilicate glass (BPSG) thin films produced in a multiple-wafer low-pressure chemical vapor deposition (LPCVD) reactor were collected to enable the development and testing of a rapid and inexpensive method for determining if films are within the desired specifications. Classification of samples into good and bad product categories was made by applying principal component analysis to the spectra. Mahalanobis distances were used as the classification metric. The highest overall percentage of correct classification of samples based upon their spectra with two-step classification was 95%. The misclassified samples were, however, within the error of the reference methods that were used in making the original classification against which the infrared (IR) classification methods were tested. The classification errors are thus just as likely to be a result of misclassification by the reference method rather than errors by the IR classification. Although reference measurements were used in this article for the original classification of the samples, these expensive and time-consuming reference methods can be eliminated simply by building classification models on samples determined to produce a product within the correct device specifications. The IR classification methods presented here hold great promise as a tool for rapid quality control of BPSG deposition. © 1997 American Vacuum Society.
[S0734-211X(97)00104-2]

I. INTRODUCTION

Borophosphosilicate glass (BPSG) is one of the dielectric materials that is widely used in the manufacturing of modern integrated circuits. At present, BPSG thin films are monitored for boron content, phosphorus content, and film thickness. The performance of BPSG films in devices can generally be predicted accurately when the boron content, phosphorus content, and film thickness of the films are known. Rapid, precise, and accurate at-line film thickness measurement tools are readily available and are widely employed. At-line measurements of film content have not been as widely adopted as have at-line film thickness measurements. Thus some manufacturers depend on time-consuming and destructive off-line measurements for boron and phosphorus content in BPSG thin films.

Fourier transform infrared (FTIR) spectroscopy, when combined with multivariate calibration, has been shown to be capable of a rapid, precise, simultaneous, and accurate determination of BPSG film thickness and content.¹⁻⁴ FTIR spectrometers are available at many microelectronic fabrication sites and can be readily employed at line to provide rapid quality control. The multivariate calibration procedure requires, however, that a relatively large set of wafers with films spanning the property ranges of interest be produced and characterized using established reference methods.

These films are used to train the multivariate calibration process. The fact that production and characterization of the training set of films are time-consuming and expensive has presented a barrier to the widespread adoption of at-line monitors that utilize multivariate calibration techniques.

The fact that a FTIR spectrum can be used to simultaneously measure the film content and film thickness with high precision suggests that very small changes in the film properties are reflected in meaningful changes in the spectral data. This spectral sensitivity to small changes in the films implies that a method might be developed using FTIR spectral data to monitor small changes in the film properties without developing a quantitative model for content and thickness. Spectral monitoring of the films can then be reduced to a classification problem. The required calibration process could be achieved by collecting spectral data from a set of "acceptable" or "good" films, i.e., films that vary due to normal process variation but remain within the process specifications. The spectral data would form a "spectral envelope" in the multidimensional data space that defines acceptable materials. A rapidly obtained infrared spectrum of an unknown film can then be compared to this envelope. If the spectral data from the unknown film is outside the envelope, then the unknown film would be classified as "unacceptable" or "bad," and a potentially faulty product would be detected.

Several well-documented methods, such as linear dis-

^{a)}Electronic mail: niemczyk@unm.edu

criminant analysis (LDA), the K -nearest neighbor (KNN) method, the Bayesian method (probability density function), and soft independent modeling of class analogy (SIMCA), have been used to classify samples based on spectral data. LDA requires a large number of training samples homogeneously collected from each class and uses a hyperplane to separate the classes.⁵ This method is not very useful for a system such as that encountered here, where we desire to build classification models using only a single class of "good" product samples. The KNN method is less dependent on homogeneity and distribution of the sampling within each class, but it does not work well if the number of training samples in each class is not approximately the same.⁶ The Bayesian method and SIMCA are both modeling approaches and are the most attractive methods for many chemical applications. The Bayesian method, however, requires a very large set of training samples to calculate the probability density function that describes the distribution of each sample class.⁷ SIMCA uses principal component analysis combined with an F test to define a hyperbox for each class,^{8,9} which does not describe the class distribution very well and is computationally complicated. Therefore, none of the established techniques is ideally suited to the classification of thin films based only on a set of good data.

The Mahalanobis distance metric, introduced by Mahalanobis in 1936,¹⁰ is a sensitive method of classifying multidimensional data. It takes the variance-covariance matrix of each training class into account and describes each class using a hyperellipsoid whose boundary is defined by the standard deviation away from the class centroid. Mark and Tunnell^{11,12} and Shah and Gemperline¹³ have reported the qualitative identification of raw materials by near infrared (NIR) spectroscopy using a Mahalanobis distance classification technique. In this article, the extension of this classification method to the fault detection of BPSG thin films using infrared absorption spectroscopy is discussed. Principal component analysis (PCA) is first applied to the data to reduce the dimensionality of the spectral data and to identify those features that change in the spectral data. The Mahalanobis distance metric is then applied to the principal components (scores) of the PCA model to develop the classification metric. When applied to a large set of BPSG thin films, we show that the combination of FTIR spectral measurements and the Mahalanobis distance classification can produce a sensitive classification technique.

II. THEORY

Principal component analysis converts the original variables to mutually orthogonal principal components (PCs) that define a reduced spectral space containing all the variations in the original spectral data that are important for classification. The PCA model can be described

$$\mathbf{A}_{(n,m)} = \bar{\mathbf{A}}_{(n,m)} + \mathbf{T}_{(n,k)}\mathbf{L}_{(k,m)} + \mathbf{E}_{(n,m)}, \quad (k \ll m), \quad (1)$$

where \mathbf{A} is the spectral data matrix with n samples and m intensities, $\bar{\mathbf{A}}$ is the matrix with a mean spectral vector in each row, \mathbf{T} is the score matrix of the first k PCs, \mathbf{L} is the

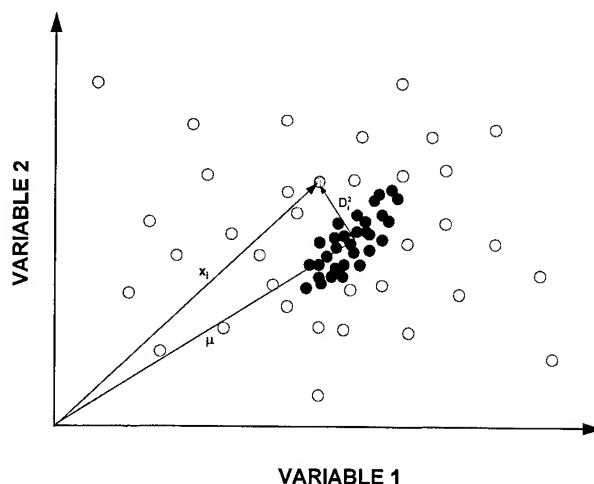


FIG. 1. A two-dimensional depiction of a Mahalanobis distance classification. The closed circles represent the samples in one class and open circles represent the second class. The shaded circle represents an unknown sample to be classified.

matrix with the first k loading vectors in rows, and \mathbf{E} is the residual matrix representing the portions of the spectral data not incorporated in the PCA classification model. Thus matrix \mathbf{A} can be represented by matrix \mathbf{T} . If the original variables are highly correlated, nearly all of the sample specific variations in the original data can be modeled by the first few PCs. In this study, the optimal number of PCs, k , was determined based on an F test of the predicted error sum of squares (PRESS) for models with fewer PCs relative to the model yielding minimum PRESS.⁷

The Mahalanobis distance is a measure of the distance from an individual point to the centroid of a population in a multidimensional space based on the assumption of a multivariate normal distribution $N(\mu, \Sigma)$ for the population. The Mahalanobis distance is expressed by

$$D_i^2 = (\mathbf{x}_i - \mu)' \Sigma^{-1} (\mathbf{x}_i - \mu), \quad (2)$$

where D_i^2 is the Mahalanobis distance, \mathbf{x}_i is a vector representing a sample, μ is a vector representing the centroid of the population, and Σ is the variance-covariance matrix of the population. In the situation described in Fig. 1, the well-defined training class can then be described by an ellipsoid with the boundary defined proportional to the standard deviation from the class centroid.

In practice, the estimated mean vector $\bar{\mathbf{x}}$ and variance-covariance matrix \mathbf{S} from a sample size of n are used instead of μ and Σ . Since all the PCs are orthogonal and the PC scores for each PC, t_{ij} ($i=1,2,\dots,n$; $j=1,2,\dots,k$), are mean centered, the estimated variance-covariance matrix \mathbf{S} calculated from the PC scores becomes a diagonal matrix with the diagonal terms given by

$$s_{jj} = \sum_{i=1}^n (t_{ij})^2 / (n-1). \quad (3)$$

The modified Mahalanobis distance for the i th sample represented by \mathbf{t}_i , a vector of k scores, can then be expressed by

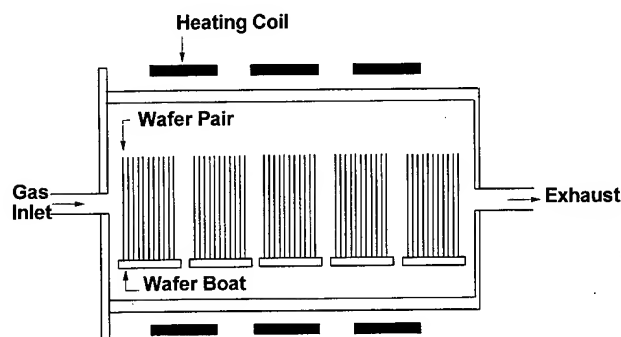


FIG. 2. Diagram of the LPCVD apparatus used to produce the 108 BPSG films.

$$D_i^2 = \mathbf{t}_i \mathbf{S}^{-1} \mathbf{t}_i' \quad (4)$$

Equation (4) can be rewritten as

$$D_i^2 = \sum_{j=1}^k (t_{ij}^2 / s_{jj}). \quad (5)$$

An unknown sample will be classified as a good or acceptable sample if its normalized Mahalanobis distance is less than unity, i.e., within the range of the original good samples used in building the classification model. In contrast to the Euclidean distance metric, the Mahalanobis distance is scaled in each dimension by the standard deviation of the scores in that dimension. Therefore, the Mahalanobis distance is given equal weight in each PC dimension and is not necessarily dominated by the early PC dimensions.

III. EXPERIMENT

The 108 BPSG thin films on 150-mm-diam undoped silicon wafers used in this study were prepared at Intel Corporation (Santa Clara, CA) in the low pressure chemical vapor deposition batch reactor depicted in Fig. 2. The reactor is capable of holding five boats of wafers with boat 1 nearest the gas inlet port and boat 5 nearest the exit (exhaust) port. Each boat holds 22 wafers as 11 back-to-back pairs. The first and the last slots of the reactor were loaded with dummy wafers thus leaving 108 test wafers. The target of the deposition process was 2.0 wt % boron, 6.7 wt % phosphorus, and 1.0 μm film thickness. The process was carried out using silane/phosphine/trimethylborane chemistry at 390 °C and 280 mTorr for 280 min. A thin capping layer of 100–200 Å undoped oxide was deposited at the end of the process to minimize the formation of boric acid crystals.

The reference measurements of the as-deposited, unannealed films were made at the center of all the films using a univariate spectral peak ratio FTIR method for boron, x-ray fluorescence (XRF) for phosphorus, and a Nanospec optical interference fringe method for film thickness. Three samples were pulled from the set to study the time-dependent effects on boron due to exposure to ambient air. Samples were defined as good if they were within $\pm 5\%$ of the target values for boron and phosphorus and within $\pm 8\%$ of the 1.0 μm target for film thickness.

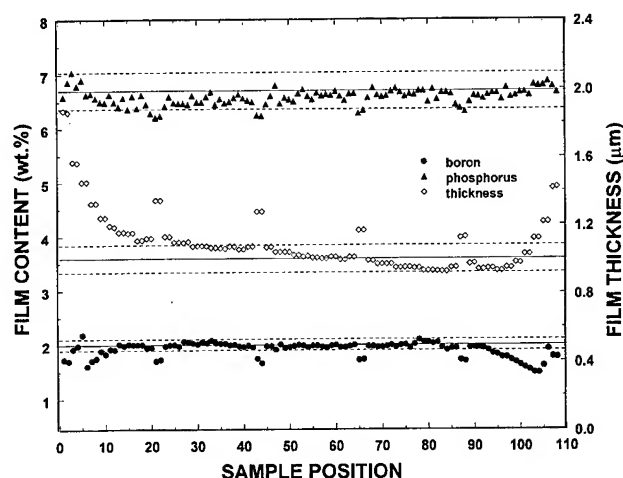


FIG. 3. Reference property determinations for the 108 BPSG films vs sample position in the LPCVD reactor. The solid horizontal lines are the property target values and the dashed lines are the property specification borders.

FTIR absorbance spectra of the unannealed thin films were collected at nine points in a cross pattern at Sandia National Laboratories with a Nicolet ECO-8S FTIR spectrometer. The wafer surfaces were thoroughly wiped with isopropyl alcohol prior to performing measurements. The spectra collected at the center of the wafers were used to test the classification procedure. All spectral data were processed using multivariate calibration and classification software developed in our laboratory.^{14,15} A number of data preprocessing options was tested with the aim of improving the classification results. The spectral preprocessing methods included first derivative, second derivative, smoothing, pathlength correction, autoscaling, range scaling, variance scaling, normalizing to unit length, and normalizing to mean-centered variance. The spectral region from 400 to 1600 cm^{-1} was used in the analyses. Twenty-five randomly selected good samples from the set of 59 good sample thin films were selected to build the classification model. The remaining 34 good and all 49 bad samples were tested as unknown samples to be classified. Mean centering and PCA on the pretreated spectral data were used to simplify construction of the mathematical model. The normalized Mahalanobis distances were calculated from the scores of the first three to five principal components and the variance-covariance matrix.

IV. RESULTS AND DISCUSSION

A. BPSG test sample set

The reference values for each of the three measured properties for all of the 108 samples are shown in Fig. 3 as a function of sample position in the reactor, where position 1 is nearest the gas inlet. It is clear that the films in boat 1 (positions 1–21) and in the latter half of boat 5 (positions 99–108) have thickness values that deviate dramatically from the target value. The first and last wafers in each boat have deviations in film thickness of about $+0.2 \mu\text{m}$, which is larger

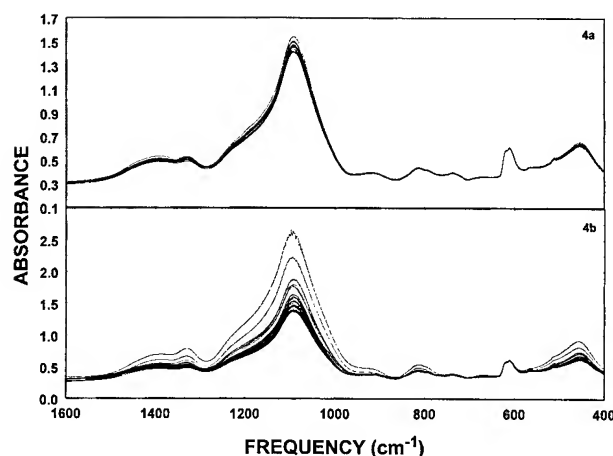


FIG. 4. Infrared spectra collected from (a) 25 randomly selected good sample films and (b) every third sample from the entire data set.

than the specified error limit. The pattern of occurrence of bad films based on the boron or phosphorus content is very similar to that of film thickness, i.e., the films at the ends of the reactor and at the ends of each boat show variations. The boron content and phosphorus content variations at the boat ends are negative, in contrast to the positive variations in film thickness. As can be seen from the data in Fig. 3, the phosphorus content is only out of the specification range for films that are also out of the specified thickness range. It is the variations in boron content that cause bad products at wafer positions 94–102.

Figure 4 contains the spectra of the 25 good samples used to develop the classification model [Fig. 4(a)] and spectra collected from every third sample, 36 total, from the entire data set. The large amount of variation in the data in Fig. 4(b) relative to that in Fig. 4(a) is indicative of the fact that IR spectral data are very sensitive to changes in the film properties. The largest changes in the data occur in regions where Si–O absorbs, 1100, 810, and 450 cm^{-1} . Changes in film thickness cause significant changes in the intensities of these bands. The variation of film thickness across the range of good samples, $1.0 \pm 0.08 \mu\text{m}$, is reflected in the variation of the strongest Si–O band, 1100 cm^{-1} , in Fig. 4(a). Note that the band at 610 cm^{-1} is invariant in both sets of spectra. This is the Si–Si phonon band, i.e., a spectral feature due to the substrate rather than to the film. The bands at ~ 1320 and $\sim 1400 \text{ cm}^{-1}$ are due to P=O and B–O, respectively. The relative intensities of these bands, when scaled for film thickness, are only approximately proportional to the boron and phosphorus content of the thin films. Previous studies have shown that a simultaneous determination of boron content, phosphorus content, and film thickness is possible if information contained across the entire 1600–400 cm^{-1} region is used.^{16,17} In fact, these prior studies have shown the IR data to be very sensitive to small changes in film properties.

B. Classification of BPSG thin films

Based on the reference values for the boron content, phosphorus content, and thickness for each film, shown in Fig. 3,

TABLE I. Classification results of the PCA/Mahalanobis distance classification using different data preprocessing methods. The test set consisted of 83 samples (34 “good” and 49 “bad”). The number of misclassified samples is noted in parentheses.

Preprocessing method	No. of PCs	% correct good	% correct bad	% correct overall
None	3	91 (3)	84 (8)	87 (11)
Smoothing	3	91 (3)	84 (8)	87 (11)
Pathlength correction	4	94 (2)	90 (5)	92 (7)
First derivative	3	97 (1)	78 (11)	86 (12)
Second derivative	5	97 (1)	47 (26)	68 (27)
Autoscaling	3	97 (1)	78 (11)	86 (12)
Range scaling	3	97 (1)	78 (11)	86 (12)
Variance scaling	3	97 (1)	78 (11)	86 (12)
Normalized to unit length	3	91 (3)	84 (8)	87 (11)
Normalized to mean-centered length	3	91 (3)	84 (8)	87 (11)

and the property specification ranges, there were 34 good samples and 49 bad samples in the test set. The spectral data were treated using a variety of preprocessing methods with the hope of identifying a method that would yield optimal classification results. In order to determine the number of principal components to use in the classification models, principal component regression (PCR) calibration was performed after preprocessing the spectral data. The PCR calibration results in the determination of an optimal number of principal components for each property based on all samples. The average optimal number of components was used for classification. In most cases, this optimal number of factors was only three factors. The preprocessing methods tested and the resulting number of principal components for each method are summarized in Table I. Also included in Table I are the results of the Mahalanobis distance classification expressed as the percentage of the samples in the test set correctly classified. The correctly classified percentages of the good, the bad, and the entire set of the test samples are listed separately.

The data in Table I show that the optimal number of principal components is three for all but two of the data preprocessing techniques. The accuracy of the classification model, when applied to good samples, was always over 90% and was relatively less sensitive to the preprocessing method applied when compared to the classification accuracy of the bad samples. Presumably the 25 good samples used to develop the classification model are representative of the good samples in the test set since these 25 samples were randomly selected.

Figure 5 indicates the positions of the misclassified samples for all of the data preprocessing options tested. A sample is flagged as misclassified in Fig. 5 even if misclassified by only one of the classification methods tested. The misclassified good samples are the same set, or subset, of three samples (samples 29, 31, and 34) for all preprocessing conditions. The classification error for all three samples could be due to the errors in the reference methods. For example, an optimistic estimate of the reference boron con-

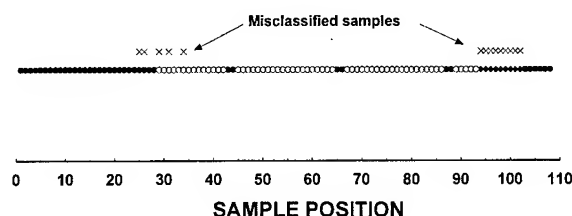


FIG. 5. Class membership vs sample position in the LPCVD reactor. (○) Good samples; (●) bad samples, film thickness out of specification; (◆) bad samples, boron content out of specification.

tent error is ± 0.1 wt %. If one takes this error into account, all three of these good samples could actually lie in the space of the bad samples (i.e., the samples may have been accurately classified, but are actually placed in the wrong category because of the poor precision of the reference methods used to place the samples in the good or bad categories).

The misclassified bad samples all come from boat 2 and boat 5. For the best model, five samples are misclassified. These five samples all have acceptable thickness and phosphorus content, but unacceptable boron content. Again, for four of these samples, the boron content as determined using the reference procedure lies within ± 0.10 wt % (the estimated error in the reference method) from the good/bad border, and the fifth is only 0.13 wt % from the border. Hence the IR misclassification might also be caused by reference determination errors. Spectral noise and classification model error might also contribute to the classification errors. However, previous results using FTIR spectral data with multivariate calibration have demonstrated that the reference measurement error of 0.10 wt % for B content dominates over spectral noise and calibration model error.

As illustrated in Fig. 4, the spectral changes due to film thickness variations are much more dramatic than are changes due to variations in boron and phosphorus content. An examination of the misclassified bad samples shows that most are samples for which the film thickness is within the range specified are acceptable, but the boron content is outside the acceptable range. This implies that the classification models are dominated by film thickness variation. The path-length correction preprocessing largely removes the effects of thickness variation, making the classification model more sensitive to changes in boron content, resulting in a classification model with the highest overall classification accuracy of the spectral preprocessing methods tested.

The large changes in the spectral data due to changes in film thickness suggest the use of a hybrid calibration/classification model. The hybrid model employs a quantitative film thickness calibration as an initial screen followed by a Mahalanobis distance classification of those samples with acceptable thickness values. The partial least-squares (PLS) algorithm was used to build a film thickness calibration model using the spectral data collected from the entire set of samples and their corresponding thickness reference measurements.^{14,18} The spectral data were baseline corrected prior to calibration. The first step of the two-step classifica-

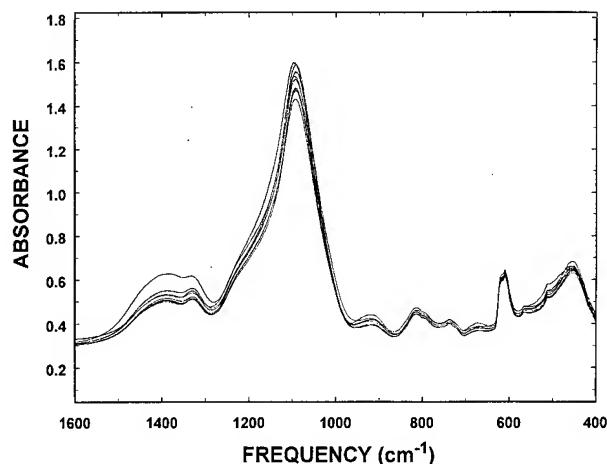


FIG. 6. Spectra collected at nine different points on a typical good sample.

tion method used the cross-validated predicted thickness values of the 83 test samples to classify samples into good and bad categories based on the film thickness values contained in the specified thickness range (e.g., $1.0 \pm 0.08 \mu\text{m}$). Those samples with thicknesses outside the specified range were immediately classified as unacceptable, i.e., bad, and not subjected to further classification. The remaining samples included all good samples and nine bad samples. All nine of these bad samples had the reference boron content outside the specified range required for good samples.

These nine bad samples and 34 good test samples were then subjected to classification using the Mahalanobis distance model built from the randomly selected 25 good samples. Of the 34 good samples, the classifier identified all but one sample correctly. This sample, position 29, is right at the boundary of the good samples and bad samples in boat 2, and is misclassified in 8 of the 10 cases listed in Table I. This result gives a 97% classification accuracy for the good test samples. When the Mahalanobis distance model was applied to the nine bad samples, all but three were correctly classified. Again, the three misclassified samples were at the boundary of the good and bad samples. This is an improvement by two samples relative to the best result achieved without film thickness screening. This two-step method improves the overall accuracy from 92% to 95%.

C. Spatial analysis of spectra over the wafers

Figure 6 shows infrared absorption spectra collected at nine different spots on a typical good sample. From Fig. 6, it can be seen that the intensities of the band centered near 1100 cm^{-1} vary from spot to spot. The band at 1100 cm^{-1} is the Si-O stretch and is highly correlated with film thickness.¹⁷ The spectral variations indicate a film thickness variation of approximately 200 \AA across the wafer based upon the PLS thickness model predictions. The spectral intensities of other spectral features in the nine spectra are more consistent with the exception of some features in the spectrum collected from the top of the wafer. In this spectrum, the bands at 1400 cm^{-1} (B-O), 1320 cm^{-1} (P=O),

920 cm^{-1} (B–O–Si), and 530–450 cm^{-1} (Si–O) have significantly higher intensities than those of the other eight spectra. Using the PLS model predictions for thickness and B, these spectral data indicate the film is thickest at the top of the wafer, and the boron content is 50% higher than at the center of the wafer. The boron content near the side edges is also higher than it is in the center region of the wafer. The phosphorus content is relatively constant throughout the entire wafer. This pattern of high boron content at the top and sides of the wafer is consistent across the test set with the exception of one or two wafers at the very ends of the reactor.

These spatial analysis results suggest that a rapid IR classification method would be able to identify portions of each wafer that are within or outside specification. Clearly an analysis method that analyzes the composition integrated over the entire wafer or a method that is so slow that it permits only a single analysis point will not be able to identify portions of a wafer that are outside specifications. Thus a rapid IR classifying method could be used to classify the wafer at each microchip location to provide potentially fewer individual chip errors and higher classification specificity.

V. CONCLUSION

Infrared spectral measurements, when combined with chemometric data interpretation, were shown to be capable of producing very accurate and precise determination of thin dielectric film properties. The method requires, however, that a carefully designed calibration set of samples be produced and characterized using reference determination methods. The production and characterization of the calibration samples is time-consuming and expensive, a fact that has limited the degree to which the method has been adopted in the microelectronics industry. The FTIR/PCA/Mahalanobis distance classification procedure presented here eliminates the complex calibration process, but retains the inherent precision and accuracy of the IR data. The classification model training is carried out by collecting IR data from good samples, and it requires no reference determinations. It is important that a large enough set of good samples be used in the training process to fully characterize the tolerable process variations.

The samples used for these studies, and those generally employed when IR monitoring of thin film production is employed, are films deposited on monitor wafers. Precise and accurate determination of dielectric thin film properties on real product wafers can be made if IR emission¹⁹ or IR reflection²⁰ data are used as the basis of the calibration. The IR/PCA/Mahalanobis distance classification could be applied to such data equally well as to absorption data. If emission or reflection data were used, the samples used in the training could be an actual good product, and the model used to classify subsequently a produced product. Such a procedure would completely eliminate the need for monitor wafers.

In all cases where the measurement precision was carefully examined, IR/chemometric determination of thin-film

dielectric properties was limited by the precision of the reference determinations used in the calibration process.^{18–20} This points out the stability of modern IR spectrometers and the inherent sensitivity of IR data. The data used for the classification process considered here are no different. The fact that a perfect classification was not achieved is most likely caused by the lack of precision of the reference determinations. The misclassified samples were all samples in which the boron content was very close, within the precision estimate for the reference method, to the good/bad border for this film property. Better boron reference determinations would likely have resulted in fewer misclassified samples.

Although not found to aid in the classification of the samples tested in this study, the film variance data showed a strong correlation to the actual classification of the samples. It is another piece of data that could aid in the classification of future samples, especially if acceptability specifications for film variability across a wafer were known.

The IR classification method described here is sufficiently rapid that classifications could be made at multiple locations on the wafers. In fact, the IR classification could be mapped on the monitor wafers to the location of individual microchips on the wafer for minimizing lost product that may only represent a portion of the wafer that is out of specification.

ACKNOWLEDGMENTS

This work was performed in part at Sandia National Laboratories and was supported by the U.S. Department of Energy under Contract No. DE-AC04-94AL85000 and by the Semiconductor Research Corporation.

- ¹B. Wangmanerat, J. A. McGuire, T. M. Niemczyk, D. M. Haaland, and J. H. Linn, *Appl. Spectrosc.* **46**, 340 (1992).
- ²J. E. Franke, L. Zhang, T. M. Niemczyk, D. M. Haaland, and J. H. Linn, *J. Electrochem. Soc.* **140**, 1425 (1993).
- ³J. E. Franke, T. M. Niemczyk, and D. M. Haaland, *Spectrochim. Acta A* **50**, 1687 (1994).
- ⁴T. M. Niemczyk, J. E. Franke, S. Zhang, and D. M. Haaland, *Mater. Res. Soc. Symp. Proc.* **341**, 119 (1994).
- ⁵D. Coomans, D. L. Massart, and L. Kaufman, *Anal. Chim. Acta* **112**, 97 (1979).
- ⁶D. Coomans and D. L. Massart, *Anal. Chim. Acta* **138**, 15 (1982).
- ⁷T. W. Anderson, *An Introduction to Multivariate Statistical Analysis* (Wiley, New York, 1958).
- ⁸S. Wold, *J. Pattern Recog.* **8**, 127 (1976).
- ⁹C. Albano, W. Dunn III, U. Edlund, E. Johansson, B. Norden, M. Sjöström, and S. Wold, *Anal. Chim. Acta* **103**, 429 (1978).
- ¹⁰P. C. Mahalanobis, *Proc. Natl. Inst. Sci. India* **2**, 49 (1936).
- ¹¹H. L. Mark, *Anal. Chem.* **59**, 790 (1987).
- ¹²H. L. Mark and D. Tunnell, *Anal. Chem.* **57**, 1449 (1985).
- ¹³N. K. Shah and P. J. Gemperline, *Anal. Chem.* **62**, 465 (1990).
- ¹⁴D. M. Haaland and E. V. Thomas, *Anal. Chem.* **60**, 1193 (1988).
- ¹⁵M. K. Alam, S. L. Stanton, and G. A. Hebner, *Spectroscopy* **9**, 30 (1994).
- ¹⁶D. M. Haaland, *Anal. Chem.* **60**, 1208 (1988).
- ¹⁷I. S. Adhihetty, J. A. McGuire, B. Wangmanerat, T. M. Niemczyk, and D. M. Haaland, *Anal. Chem.* **63**, 2329 (1991).
- ¹⁸J. E. Franke, L. Zhang, T. M. Niemczyk, D. M. Haaland, and K. J. Radigan, *J. Vac. Sci. Technol. A* **13**, 1959 (1995).
- ¹⁹S. Zhang, Ph.D. dissertation, University of New Mexico, 1996.
- ²⁰L. Zhang, Ph.D. dissertation, University of New Mexico, 1994.

Highly preferred (111) texture aluminum-copper films formed with argon plasma treatment of the titanium underlayer and their electromigration endurance as interconnects

Kazuyoshi Kamoshida^{a)} and Yasuyuki Ito

NTT System Electronics Laboratories, 3-1 Morinosato Wakamiya, Atsugi-shi, Kanagawa 243-01, Japan

(Received 7 February 1997; accepted 16 May 1997)

A new fabrication technology for producing hypertextured aluminum-copper (Al-Cu) films has been developed. The key feature of this technology is the argon (Ar) plasma treatment of the titanium (Ti) underlayer. The full width at half-maximum values of the Al (111) rocking curve decrease with increasing Ti etched thickness. That is, the texture of the Al-Cu film improves with increasing Ti etched thickness. The Ti surface treated with Ar plasma has a high surface energy because the Al island wetting angle is small. We think this means the Ar-plasma-treated Ti surface structure is amorphouslike and rough. Consequently, highly preferred (111) texture Al-Cu films were obtained with argon plasma treatment of the Ti underlayer. The electromigration (EM) lifetime of these films is one order of magnitude higher than that of films obtained without the plasma treatment. The longer EM lifetime of the hypertextured Al lines is considered to be due to the small grain-boundary diffusivities for the bamboo microstructure. The argon plasma treatment of the titanium underlayer, combined with Al-Cu deposition, makes it possible to form highly EM-resistant multilevel interconnections without any of the drawbacks of the conventional method.

© 1997 American Vacuum Society. [S0734-211X(97)14104-X]

I. INTRODUCTION

Electromigration (EM) failure in aluminum (Al) interconnections has become an increasingly serious problem in large-scale integrated circuits.¹⁻⁴ Efforts to make metal lines less susceptible to EM have focused on creating new alloys such as Al-Si-Ti,⁵ Al-Cu-Ti,⁶ Al-Pd-Si,⁷ and Al-Sc⁸ and on using a layered structure with underlying TiN/Ti⁹ or TiN¹⁰⁻¹² barrier metals.

Microstructures strongly influence EM behavior in thin Al films. The most thorough study of microstructural effects was conducted by Vaidya and Sinha,¹³ who varied grain size, grain size distribution, and texture over a variety of processing conditions for Al-0.5%Cu. Their results produced the following relation between microstructural variables and median time to failure (MTTF):

$$\text{MTTF} \propto \frac{S}{\sigma^2} \log \left(\frac{I_{(111)}}{I_{(200)}} \right)^3, \quad (1)$$

where S is the median grain size, σ is the distribution of the Al grain sizes, and $I_{(111)}$ and $I_{(200)}$ are the x-ray intensities of the Al (111) and Al (200) peaks. Equation (1) suggests that the highly preferred (111) texture is more resistant to EM. A variety of deposition techniques^{14,15} and substrates¹⁶⁻¹⁹ have been used in attempts to improve Al film texture, and Toyoda *et al.* showed that using an amorphous Al-Ta underlayer, which has a large surface energy, produces overlying Al (111) with excellent texture.²⁰

This article investigates the effects of an Ar-plasma-treated Ti film on the overlying aluminum-copper (Al-Cu) layered interconnects. Results show that the crystallographic

orientation of Al-Cu films deposited on a treated Ti surface changes in such a way that Al-Cu/Ti layered interconnects formed on the treated Ti film exhibit a highly reliable EM performance.

II. EXPERIMENT

A. Film preparation and evaluation

The samples used to evaluate the properties of aluminum alloy (Al-0.5 wt. %Cu)/Ti films and the effect of the Ar plasma treatment of the Ti surface were fabricated as follows. A 300-nm-thick SiO₂ film was deposited by plasma chemical-vapor deposition (P-SiO₂) on 6 in. Si wafers using SiH₄+N₂O gases at a substrate temperature of 450 °C. Titanium films 5 to 100 nm thick were then deposited on the unheated P-SiO₂/Si substrate using a multichamber dc magnetron sputtering system. The base pressures before Ti deposition were below 1×10^{-9} Torr (1.3×10^{-7} Pa), and the Ar gas pressure during sputtering was 2 mTorr (0.27 Pa). The deposition rate was 0.8 nm/s. The Ti target purity was 99.9999%; the oxygen content was lower than 100 ppm. The Ti surface was etched with Ar to the Al-Cu deposition using inductive coupled plasma (ICP) etching equipment (Fig. 1). The rf power applied to the substrate was 13.56 MHz and that applied to the ICP coil was 400 kHz. The vacuum system in ICP was turbo pumped and the base pressure before etching was about 1×10^{-8} Torr (1.3×10^{-6} Pa). Etching pressure was 0.4 mTorr (0.05 Pa) at the ICP power of 205 W, and the bias power was 175 W. The Ti surface etching rate was approximately 36 nm/min.

Films of Al-Cu 10–500 nm thick were deposited on the Ar-plasma-treated Ti/P-SiO₂ structure at 100 and 300 °C using the multichamber dc magnetron sputtering system. In

^{a)}Electronic mail: kamosida@aecl.ntt.co.jp

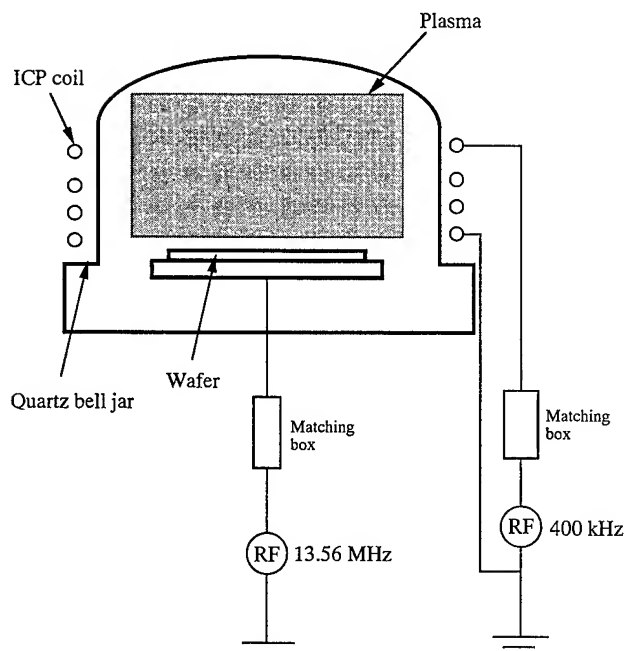


FIG. 1. Schematic diagram of the Ar-plasma-treatment apparatus.

other samples, Al-Cu films were deposited, also at 100 and 300 °C, on Ti/P-SiO₂ and directly on the P-SiO₂ layer without a Ti underlayer. An Al-Cu film was deposited on the Ar-plasma-treated Ti, and the untreated Ti substrate without breaking vacuum. The base pressure before Al-0.5 wt. %Cu deposition was below 1×10^{-9} Torr (1.3×10^{-7} Pa), and the Ar gas pressure during sputtering was 2 mTorr (0.27 Pa). The deposition rates varied from 3.3 to 19.2 nm/s.

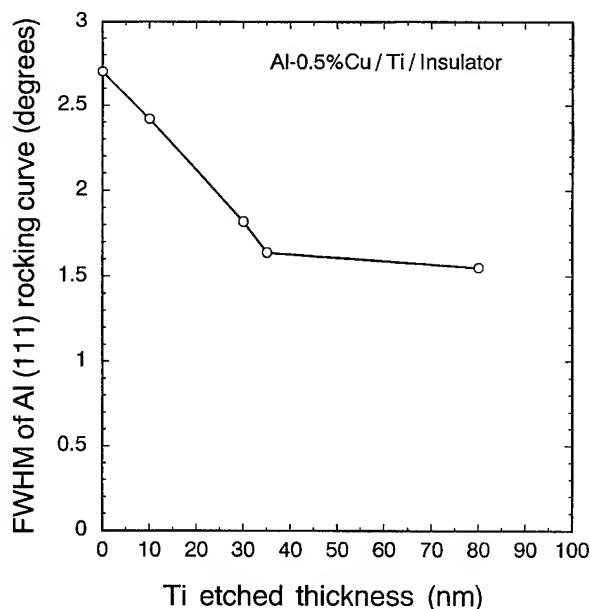


FIG. 2. FWHM values obtained from Al-Cu (500 nm)/Ti formed on P-SiO₂ as a function of Ti etched thickness. Final Ti thickness was 20 nm.

TABLE I. Average Al grain size, dispersion of the Al grain size, and FWHM values of Al films.

Ar plasma treatment	Average grain size (nm)	Standard deviation (nm)	FWHM of Al (111) rocking curve (deg)
Without	220	33	2.7 ^a
With	200	38	1.8 ^b

^aTi thickness: 20 nm.

^bTi etched thickness: 30 nm; remaining Ti thickness: 20 nm.

The degree of the (111) texture was evaluated in terms of the full width at half-maximum (FWHM) of the rocking curve for the Al (111) peak. The rocking curve was measured by x-ray diffraction (XRD), using a Cu $K\alpha$ line, by scanning at an incident angle θ while the diffraction angle 2θ was fixed for the Al (111) peak at 38.47°. The grain sizes and cross sections of the Al-Cu films were observed by cross-sectional transmission electron microscopy (XTEM), the surface morphology was observed by scanning electron microscopy (SEM), and the surface roughness of the Ti film was evaluated by atomic force microscopy (AFM). The argon concentration in the Ti films was determined by a total reflection x-ray fluorescence (TXRF) analysis. The TXRF analyzer was equipped with a tungsten anode x-ray tube. The high voltage was 30 keV at a tube current of 200 mA.

B. Electromigration test

For EM evaluation, the samples were TiN (30 nm)/Ti (10 nm)/Al-0.5%Cu (500 nm) metal lines formed on Ti (20 nm)/P-SiO₂ (300 nm) double layer films with and without Ar plasma treatment. The layered metal lines were 480 nm wide and 1 mm long. A TiN layer was deposited by reactive sputtering with a pure Ti target in 40% N₂ ambient without substrate heating. The base pressures before TiN deposition were below 1×10^{-9} Torr (1.3×10^{-7} Pa), and the Ar+N₂ gas pressure during sputtering was 3 mTorr (0.4 Pa). The deposition rate was 0.65 nm/s. The purpose of the

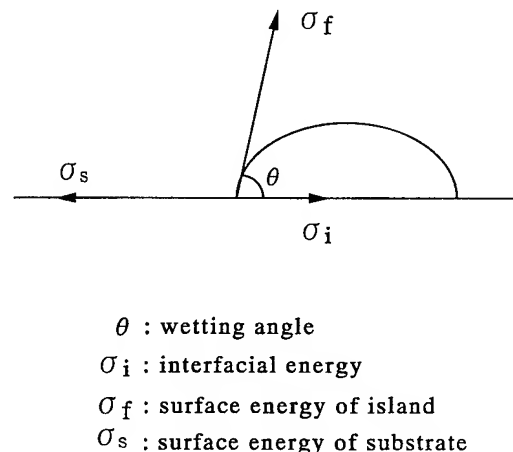


FIG. 3. Schematic illustration of the terms in Young's equation.

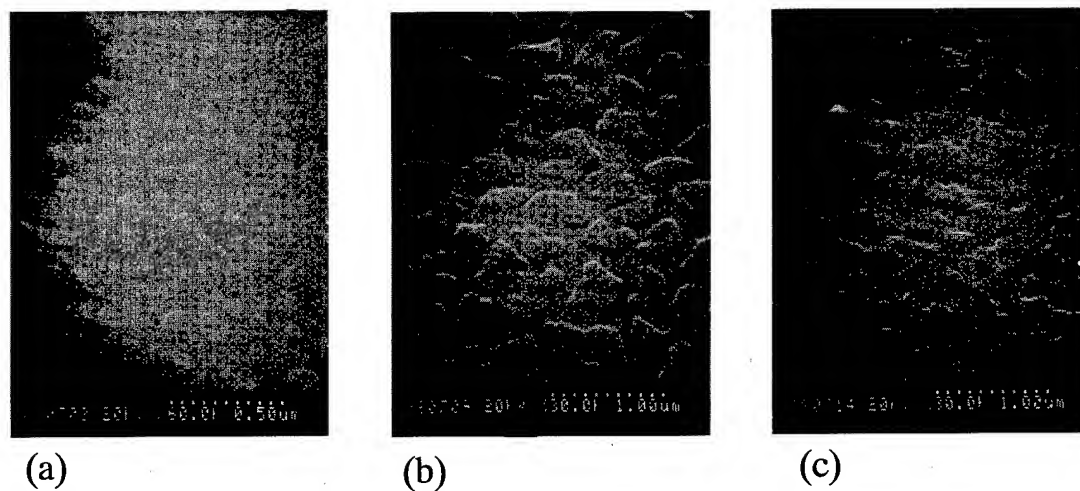


FIG. 4. SEM observation of the Al-Cu surface morphologies. Al-Cu films formed on (a) P-SiO₂ (300 nm), (b) Ti (5 nm), and (c) an Ar-plasma-treated Ti underlayer (Ti etched thickness: 25 nm; remaining Ti thickness: 5 nm). Al-Cu deposition temperature was 300 °C. Al-Cu thickness was 10 nm.

top layer of TiN was to serve as the antireflective coating layer for patterning of submicrometer-level interconnects by photolithography. That of the Ti layer was to prevent formation of AlN during the deposition of TiN using Ar+N₂ plasma on the Al-Cu layer.^{21,22} The Al-Cu film deposition temperature was 100 °C and the Ti thickness after Ar plasma etching was 20 nm.

For the EM tests, fully processed wafers were used with four-terminal stripes. After deposition, the multilayered structures [TiN (30 nm)/Ti (10 nm)/Al-0.5%Cu (500 nm)/Ti (20 nm)/P-SiO₂ (300 nm)] were patterned by lithography and reactive ion etching in order to produce interconnects. The EM-test samples were unpassivated, and EM-test wafers were subjected to a continuous dc current density of 1×10^7 A/cm² at a temperature of 200 °C using a resistively heated sample holder. The failure criterion was a 10% increase from the initial metal line resistance.

III. RESULTS AND DISCUSSION

A. Effects of Ar-plasma-treated Ti underlayer on Al film properties

The FWHM value of Al (111) rocking curve for the Al-0.5 wt. %Cu films is plotted in Fig. 2 against the Ti etched thickness. In this case, the remaining Ti thickness was 20 nm, the metal structure was Al-0.5%Cu/Ti/P-SiO₂, and the Al-Cu was deposited at 100 °C to a thickness of 500 nm. The FWHM values decrease with increasing Ti etched thickness. That is, the texture of the Al-Cu films improved with increasing Ti etched thickness. The FWHM value of Al (111) varied from about 2.7°, without plasma treatment, to 1.8° when the Ti etched thickness was 30 nm and to 1.5° when the Ti etched thickness was 80 nm. These results mean Al-Cu film with the highly preferred (111) texture were obtained with the Ar-plasma-treated Ti samples.

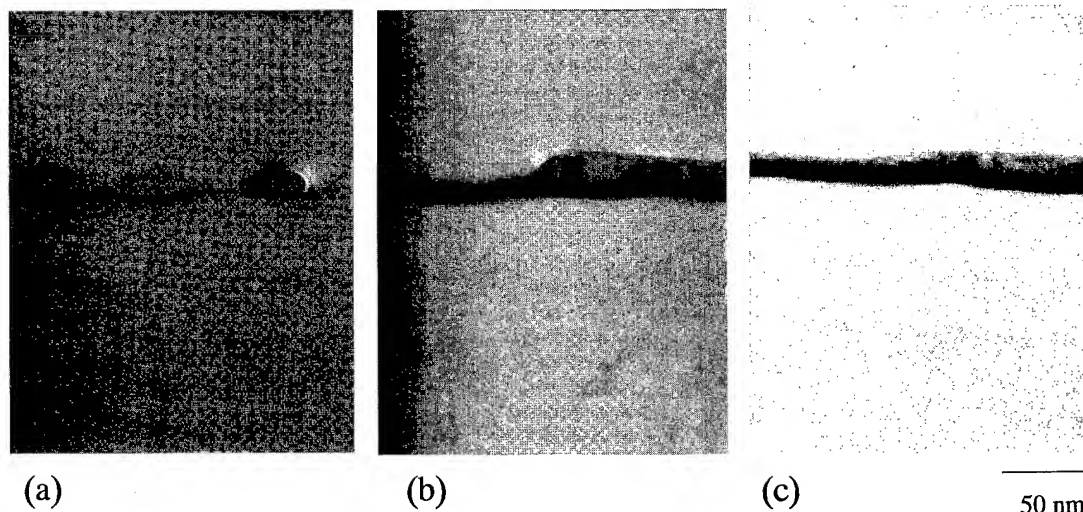


FIG. 5. Cross-sectional TEM view of the Al-Cu films formed on (a) P-SiO₂ (300 nm), (b) Ti (5 nm), and (c) an Ar-plasma-treated Ti underlayer (Ti etched thickness: 25 nm; remaining Ti thickness: 5 nm). Al-Cu deposition temperature was 300 °C. Al-Cu thickness was 10 nm.

TABLE II. Al wetting angle.

Underlayer	Wetting angle (deg)
P-SiO ₂	144
Ti (without Ar treatment)	36 ^a
Ti (with Ar treatment)	7 ^b

^aTi thickness: 5 nm; Al thickness: 10 nm.

^bTi etched thickness: 25 nm; remaining Ti thickness: 5 nm; Al thickness: 10 nm.

We measured the average Al grain size and the distribution of the Al grain size. The average Al grain size, the distribution of the Al grain size, and the FWHM values of the Al films are summarized in Table I. The average Al grain sizes and distributions are almost the same with and without the treatment, but the FWHM values of the Al (111) rocking curve are different: 2.7° with no treatment and 1.8° with treatment.

B. Initial growth of Al-Cu film

In classical theory, island shape (Fig. 3) is determined by Young's equation

$$\cos \theta = \frac{\sigma_s - \sigma_i}{\sigma_f}, \quad (2)$$

where θ is the wetting angle of the island, σ_s and σ_f are the respective surface energies of the substrate and the island, and σ_i is the interfacial energy between the island and the substrate. Toyoda *et al.* suggested that when $\cos \theta$ is nearly 1 ($\theta \approx 0$) the film grows two dimensionally (layer growth).²⁰ In this layer-growth mode, it is expected that two-dimensional islands form the most stable (111) closely packed configuration, especially for face centered cubic structure (fcc) metals such as Al and Cu. Mitsuzuka reported that Al films on ultrathin amorphouslike metal buffer layers have a good texture but that Al films on thick buffer layers have poor texture even though the surface energy of the ultrathin and thick buffer layers is the same.¹⁹ This suggests that an amorphous state of the underlayer is very important for improving the texture. Hasunuma *et al.* suggested that an amorphous substrate will enhance layer growth because there is no specific crystal-lattice-matching restriction between a film and substrate that obstructs stable (111) two-dimensional island formation.²³ The (111) texture of the film is improved by the layer-growth mode according to Eq. (2), high σ_s and low σ_i are preferable for obtaining a small wetting angle. These

TABLE III. Titanium surface roughness.

Ar plasma treatment	Roughness (rms: nm)
Without	0.26 ^a
With	1.22 ^b

^aTi thickness: 20 nm.

^bTi etched thickness: 30 nm; remaining Ti thickness: 20 nm.

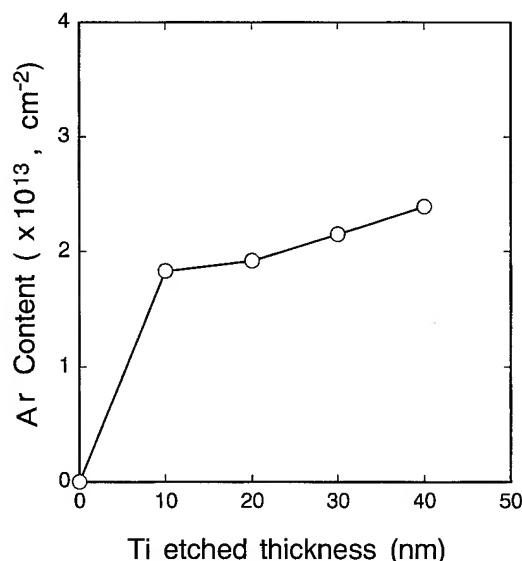


FIG. 6. Argon content determined by TXRF analysis vs Ti etched thickness.

previous results indicated that both a small wetting angle (high surface energy or low interfacial energy) and an amorphous state are necessary for improving the texture.

Figure 4 shows surface morphologies observed by SEM when Al-Cu films were grown on (a) P-SiO₂, (b) untreated Ti, and (c) treated Ti substrates. The Al-Cu islands formed on P-SiO₂ [Fig. 4(a)]. The Al-Cu film grown on untreated Ti is not uniform and has a semispherical surface [Fig. 4(b)]. On the other hand, the Al-Cu film grown on the treated Ti is smooth and uniform [Fig. 4(c)]. Figure 5 shows XTEM images of the initial Al-Cu film growth in Fig. 4. The island-type growth is apparent on the P-SiO₂ substrate [Fig. 5(a)], and one can see the slightly rugged surface the Al-Cu films growth on the untreated Ti [Fig. 5(b)]. In contrast, it is clear that continuous Al-Cu films were grown on the treated Ti [Fig. 5(c)].

We therefore next measured the wetting angle of the Al island so that we could compare the surface energy of the substrate between samples with and without Ar treatment. The substrate dependence of the wetting angle is summarized in Table II: Al-island wetting angles were 144° on P-SiO₂, 36° on Ti, and 7° on Ar-plasma-treated Ti. These results suggest that surface energy of Ar-plasma-treated Ti samples is larger than that of the untreated Ti surface and the P-SiO₂ surface. The present study thus indicates that the Ar-plasma-treated Ti surface has a high surface energy because the Al-island wetting angle is small.

Figure 6 shows the relation between the argon concentration and the Ti etched thickness. The argon content of the Ti films increased with etched thickness, from about $1.8 \times 10^{13} \text{ cm}^{-2}$ at 10 nm to about $2.2 \times 10^{13} \text{ cm}^{-2}$ at 40 nm. This suggests that the Ti surface structure consists of a mixture of Ti and Ar. From these results, we think that the Ar-plasma-treated Ti surface structure is amorphouslike.

Figure 7 shows the Ti surface morphologies observed by AFM for untreated and treated Ti. The untreated Ti surface is smooth and the treated surface is rough and has a large grain

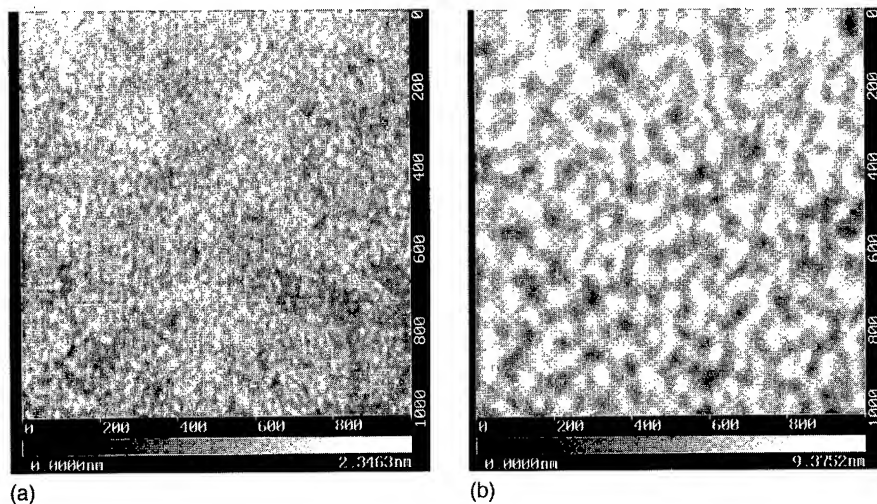


FIG. 7. AFM observation of the Ti surface morphologies. (a) Without Ar-plasma-treated Ti (Ti thickness: 20 nm), (b) an Ar-plasma-treated Ti (Ti etched thickness: 30 nm; remaining Ti thickness: 20 nm).

size. Table III lists the rms roughness values calculated from AFM images of Ti surfaces with and without Ar plasma treatment. These results suggested that the surface contained a large number of Al-Cu nucleation sites. Consequently, it is postulated that the continuous Al-Cu film (two-dimensional island formation) growth would proceed in the early stage. This possibility will be examined in future experiments.

C. EM performance in Al-Cu/Ti layered interconnects

We used XTEM to evaluate the grain boundary distribution of the Al-Cu films grown on Ti underlayers with and without the Ar plasma treatment. Figure 8 shows that there is a large difference between treated and untreated samples:

Without the treatment the grain boundaries formed at random, whereas with the treatment they are perpendicular to the substrate and bamboo microstructure.

Figure 9 shows the results of EM tests carried out on metal lines formed on samples with and without the Ar plasma treatment of the Ti underlayer. The initial resistance values of the samples at time $t=0$ were measured without (208.9Ω) and with (220.2Ω) Ar plasma treatment. With the Ar plasma treatment, the MTTF is ten times what it is without it (1.6×10^5 vs 1.6×10^4 s). The longer EM lifetime for the hypertextured Al line is considered to be due to the small grain boundary diffusivities of the bamboo microstructures: this low diffusivity suppresses formation of void/hillock pairs in the Al-Cu lines.

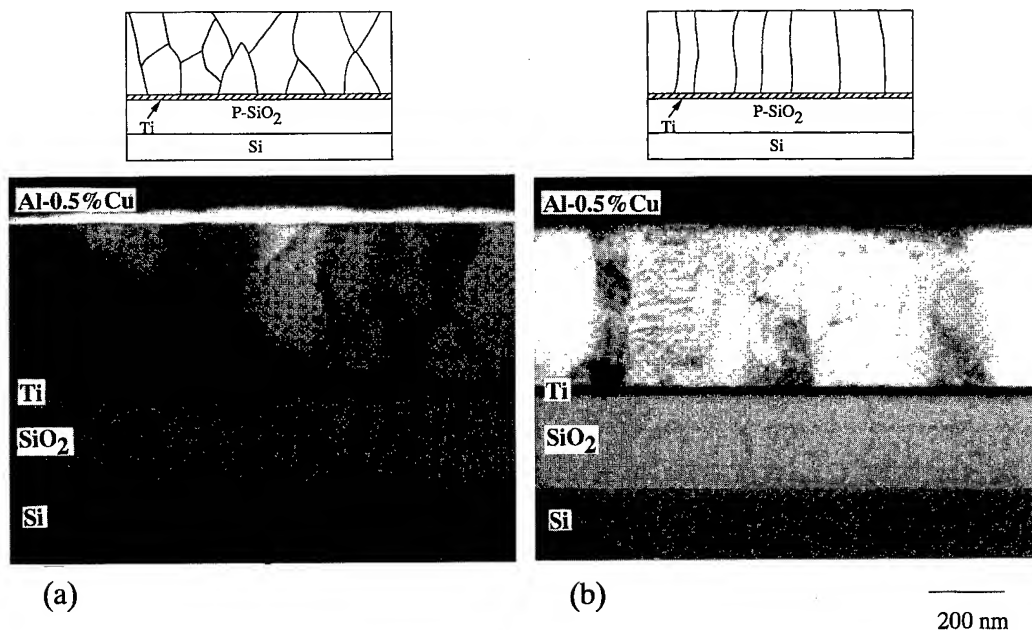


FIG. 8. Cross-sectional TEM view of 500 nm-thick Al-0.5%Cu films formed on Ti (20 nm) P-SiO₂ (300 nm) double layer films (a) without and (b) with Ar plasma treatment of the Ti underlayer.

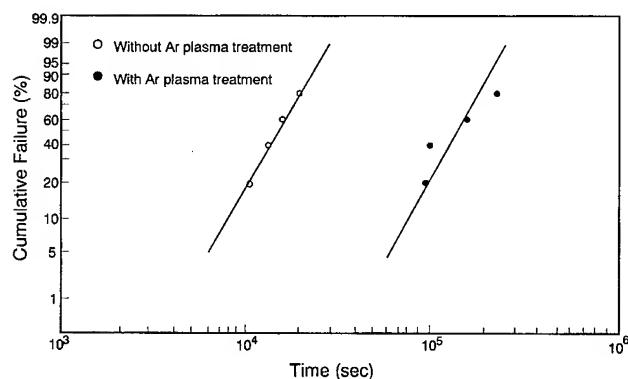


FIG. 9. EM test results. The samples were TiN (30 nm)/Ti (10 nm)/Al-0.5%Cu (500 nm) layered metal line formed on Ti (20 nm)/P-SiO₂ (300 nm) double layer films with and without Ar plasma treatment. The Al-Cu deposition temperature was 100 °C, and dc current density during the test at 200 °C was 1×10^7 A/cm².

IV. CONCLUSION

We have developed hypertextured Al-Cu film formation technology based on Ar plasma treatment of a Ti underlayer. The FWHM values decrease with increasing Ti etched thickness. That is, the texture of Al-Cu film improves with increasing Ti etched thickness. The Ar-plasma-treated Ti surface has a high surface energy because the Al island wetting angle is small. We think this means Ar-plasma-treated Ti surface structure is amorphouslike and rough. Consequently, highly preferred (111) texture Al-Cu films were obtained with argon plasma treatment of the Ti underlayer. The EM lifetime in films grown on treated underlayers is one order of magnitude higher than in films grown on untreated underlayers. This longer EM lifetime for the hypertextured Al lines is thought to be due to the small grain-boundary diffusivities for the bamboo microstructure.

The argon plasma treatment of the titanium underlayer, combined with Al-Cu deposition, makes it possible to form highly EM-resistant multilevel interconnections without any of the drawbacks of the conventional method.

ACKNOWLEDGMENTS

The authors would like to thank Drs. Masaaki Sato, Yoshinobu Arita, Katsutoshi Izumi, and Tetsushi Sakai

(NTT Electronics Technology Inc.) for their encouragement and for suggesting this study. They also thank Kazushige Uematsu (NTT Advance Technology Inc.) and Keiji Himei (NTT Electronics Technology Inc.) for the XRD analysis and for sample preparation, and they thank Masao Nagase for the AFM analysis.

¹D. B. Fraser, *VLSI Technology*, edited by S. M. Sze (McGraw-Hill, New York, 1983), p. 347.

²D. Pramanik and A. N. Saxena, *Solid State Technol.* **26**, 131 (1983).

³P. B. Ghatge, *Solid State Technol.* **26**, 113 (1983).

⁴P. S. Ho and T. Kwok, *Rep. Prog. Phys.* **52**, 301 (1989).

⁵J. M. Towner, A. G. Dirks, and T. T. Tien, *Proceedings of the 24th International Reliability Physics Symposium*, Anaheim, CA, 1986 (unpublished), p. 7.

⁶T. Hosoda, H. Yagi, and T. Tsuchikawa, *Proceedings of the 27th International Reliability Physics Symposium*, Phoenix, AZ, 1989 (unpublished), p. 202.

⁷Y. Koubuchi, J. Ohnuki, S. Fukada, and M. Suwa, *IEEE Trans. Electron Devices* **37**, 947 (1990).

⁸S. Ogawa and H. Nishimura, *Technical Digest, International Electron Devices Meeting*, Washington, D.C., 1991 (unpublished), p. 277.

⁹M. Sekiguti, K. Sawada, M. Fukumoto, and T. Kouzaki, *J. Vac. Sci. Technol. B* **12**, 2992 (1994).

¹⁰J. S. Byun, K. G. Rha, J. J. Kim, W. S. Kim, H. N. Kim, H. S. Cho, and H. J. Kim, *J. Appl. Phys.* **73**, 1719 (1995).

¹¹Y. Koubuchi, S. Ishida, M. Sahara, Y. Tanigaki, T. Kato, J. Onuki, and M. Suwa, *J. Vac. Sci. Technol. B* **10**, 143 (1992).

¹²H. Onoda, M. Kageyama, and K. Hashimoto, *J. Appl. Phys.* **77**, 885 (1995).

¹³S. Vaidya and A. K. Sinha, *Thin Solid Films* **75**, 253 (1981).

¹⁴Y. H. Park, P. Roessle, E. Majewski, and J. F. Smith, *J. Vac. Sci. Technol. A* **3**, 2308 (1985).

¹⁵D. B. Knorr and T.-M. Lu, *Appl. Phys. Lett.* **54**, 2210 (1985).

¹⁶K. Hashimoto and H. Onoda, *Appl. Phys. Lett.* **10**, 120 (1989).

¹⁷M. Kageyama, K. Hashimoto, and H. Onoda, *Proceeding of the 29th International Reliability Physics Symposium*, Las Vegas, NV, 1991 (unpublished), p. 97.

¹⁸H. Shibata, M. Murota, and K. Hashimoto, *Jpn. J. Appl. Phys.* **32**, 4479 (1993).

¹⁹T. Mitsuzuka, *Jpn. J. Appl. Phys.* **31**, L1280 (1992).

²⁰H. Toyoda, T. Kawanoue, M. Hasunuma, H. Kaneko, and M. Miyauchi, *Proceeding of the 32nd International Reliability Physics Symposium*, San Jose, CA, 1994 (unpublished), p. 178.

²¹H. Onoda, M. Kageyama, Y. Tatara, and Y. Fukuda, *IEEE Trans. Electron Devices* **40**, 1614 (1993).

²²Y. Inoue, S. Tanimoto, K. Tsujimura, T. Yamashita, Y. Ibara, Y. Yamashita, and K. Yoneda, *J. Electrochem. Soc.* **141**, 1056 (1994).

²³M. Hasunuma, H. Toyoda, T. Kawanoue, S. Ito, H. Kaneko, and M. Miyauchi, *Mater. Res. Soc. Symp. Proc.* **391**, 335 (1995).

Surface nitridation of silicon dioxide with a high density nitrogen plasma

R. Kraft,^{a)} T. P. Schneider, W. W. Dostalík, and S. Hattangady

Semiconductor Process and Device Center, Texas Instruments Inc., Dallas, Texas 75243

(Received 20 February 1997; accepted 16 May 1997)

A high density nitrogen plasma generated with a helicon plasma source has been used to incorporate approximately 15 at. % nitrogen into the top 0.5 nm of a silicon dioxide layer. The surface nitridation was accomplished in 10 s with a high flux of low energy ions which were extracted from the high density nitrogen plasma and accelerated in the plasma sheath towards the surface an electrically floating silicon dioxide surface. A rf compensated Langmuir probe was used to measure the nitrogen ion energy and ion current density as a function of the nitrogen pressure and source power. The nitrogen ion energy, ion current density, and exposure time determine the nitrogen range and dose into the silicon dioxide surface. This process may be advantageous for nitriding the gate oxide in advanced complementary metal-oxide semiconductor process flows. © 1997 American Vacuum Society. [S0734-211X(97)14204-4]

I. INTRODUCTION

Plasma nitridation is a useful technique for modifying the physical characteristics of materials.^{1,2} For example, in the complementary metal-oxide semiconductor (CMOS) process flow it may be desirable to nitride the surface of a material to change its electrical properties, oxidation resistance, etch resistance, or diffusion barrier properties. When the nitridation process described in this article was applied to a 4.0 nm gate oxide in a 0.18 μm CMOS process flow,^{3,4} boron penetration from the doped polysilicon gate was suppressed. This was accomplished with (1) no degradation in n -channel or p -channel mobility, (2) drive currents which were equivalent to, or exceed that of control oxide, (3) little flat band voltage shift, and (4) no significant variation in midgap interface trap density from that of control oxide.

The attractive feature of plasma nitridation for CMOS applications is that high nitrogen concentrations can be confined very close to the top surface of the nitrided material. Compared with typical high thermal budget nitridation processes, such as with N_2O or NO ,^{5,6} the plasma nitridation process incorporates nitrogen much closer to the surface and does not put any nitrogen at the oxide/silicon interface. Compared with ion implantation nitridation, plasma nitridation can be performed at much lower ion energies and much higher ion current densities. In particular, a low pressure, high density nitrogen plasma generated in an inductively coupled, electron cyclotron, or Helicon source can deliver ion current densities of 5 mA/cm^2 and ion energies of 10 eV over large areas such as a 300 mm wafer. High nitrogen concentration can be attained in a short period of time, thus making the plasma nitridation process suitable for high volume semiconductor manufacturing.

II. THEORY

In the plasma nitridation process we describe in this article, the nitrided substrate was allowed to float in potential

to attain low nitrogen ion energy to the substrate surface. This resulted in a shallow nitrogen incorporation range. When the substrate is unbiased, it will attain a potential known as the floating potential V_f , while the bulk plasma will attain a more positive potential known as the plasma potential V_p . The positive plasma potential arises to contain the electrons, which are more mobile than the ions, and therefore maintain charge neutrality in the bulk of the plasma. The floating potential arises to balance the net electron and ion current to the substrate which must be in the steady state condition.

The ion current to the substrate⁷ is expressed as

$$J_f = 0.6N_i e \langle v_i \rangle = 0.6N_i e \sqrt{\frac{8T_e}{\pi m_i}}, \quad (1)$$

where e is the electronic charge, $\langle v_i \rangle$ is the ion random thermal velocity, m_i is the ion mass, and T_e is the electron temperature. The Bohm criterion was used in Eq. (1) to relate $\langle v_i \rangle$ to T_e . The high density Helicon plasma source used in this study was capable of generating ion densities in the 10^{10} – 10^{11} ions/ cm^3 range which resulted in ion current densities in the 1–5 mA/cm^2 range.

The ion energy to the substrate is mainly determined by the difference between the plasma and floating potential. The ions acquire this energy as they pass through the sheath region at the substrate surface. The ion energy is therefore expressed as

$$E_i \sim e(V_p - V_f). \quad (2)$$

The term $(V_p - V_f)$ can be related to electron temperature by equating the ion and electron currents at the substrate surface. Assuming a Maxwell-Boltzmann electron energy distribution, the relationship for the ion energy becomes

$$e(V_p - V_f) = \frac{T_e}{2} \ln \left(\frac{m_i}{2\pi m_e} \right) = E_i, \quad (3)$$

^{a)}Present address: Semiconductor Process and Device Center, Texas Instruments, Inc., 13536 N. Central Expressway, Dallas, TX 75243; Electronic mail: kraft@spdc.ti.com

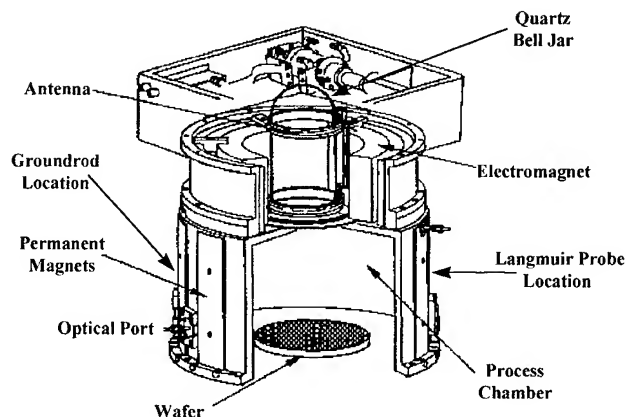


FIG. 1. Cross section of the experimental process module.

where m_e is the electron mass. The high density Helicon plasma source used in this study generated electron temperatures in the 2–3 eV range which resulted in average ion energies in the 8–14 eV range.

III. EXPERIMENTAL APPARATUS

The experiment was performed in a standard process module on a commercial etch tool.⁸ The process module (Fig. 1) consisted of (1) a Helicon^{9,10} plasma source powered by a 13.56 MHz rf generator with a maximum 3300 W output capability (the Helicon wave was excited with a Nagoya type III antenna which was located outside of the quartz bell jar), (2) a pair of electromagnets to generate the magnetic field required to support the Helicon wave in the plasma, and (3) a multipole fixed magnet reaction vessel for plasma confinement.

The controllable process conditions were nitrogen flow, chamber pressure, power to the Helicon source, magnetic field strength, and process time. An automatic matching network was used to ensure a good match between the rf generator and the Helicon source. The dc currents through the inner and outer electromagnets were adjusted to modify the magnetic field profile in the chamber and attain the best plasma uniformity at the wafer surface. The magnetic field in the bell jar region was approximately 125 G and approximately 0 G at the wafer surface. The nitrogen gas was introduced through holes in the upper plate of the chamber and the flow was regulated at 100 sccm with a molecular flow controller. The chamber pressure was controlled with a throttle valve, pressure transducer, and a turbo-molecular pump.

The substrates used in this study were 150 mm diameter (100) silicon wafers with a 40 Å thermally grown oxide layer. The wafer was loaded into the chamber by an automatic wafer handling system. The distance between the bottom of the quartz bell jar and the wafer surface was 18.7 cm. The wafer was held on a ceramic lift ring well away from any grounded surfaces in the chamber.

Time averaged plasma parameters were measured 5 cm above the wafer surface with a commercial rf compensated single Langmuir probe.¹¹ The details of the theory and op-

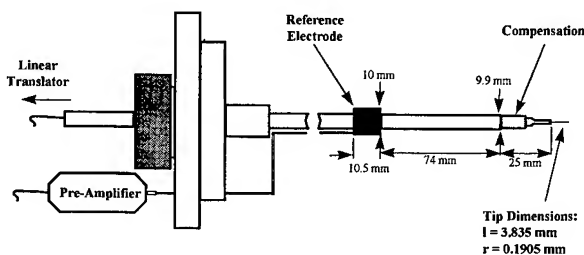


FIG. 2. Schematic of the rf compensated Langmuir probe. The high frequency compensation for plasma potential measurements and the reference electrode for low frequency compensation are noted.

eration of the Langmuir probes can be found elsewhere.¹² The plasma parameters measured included the electron temperature, T_e , ion density, N_i , ion saturation current, J_1 , floating potential, V_f , and the plasma potential, V_p . All the plasma measurements were uniform to within a few percent across the wafer surface and the reported plasma parameters corresponding to the center of the wafer.

The Langmuir probe system (Fig. 2) consisted of a tungsten tip, stainless steel high frequency compensation electrode, ceramic rod, stainless steel reference electrode, and a vacuum flange including electrical feedthroughs. The entire system was attached to a linear translator that enabled radial spatial plasma measurements over a range of 25 cm. Since the plasma in the chamber was confined from the walls by permanent magnets, we found it was necessary to install a 10 cm long tungsten ground rod into the plasma to ensure a good return current path for the Langmuir probe. The ground rod (Fig. 1) was positioned in the same plane as the Langmuir probe and was long enough to pass through the magnetic field from the permanent magnets.

IV. EXPERIMENTAL PLASMA MEASUREMENTS

Plasma measurements were made for source powers of 750–2500 W and pressures of 4–20 mT. The electron temperature and difference between the plasma and the floating potentials are shown in Figs. 3 and 4. The electron temperature was relatively insensitive to the source power but it did decrease as the pressure was raised possibly due to an increase in the number of electron-ion collisions. The difference between the plasma and floating potentials (which de-

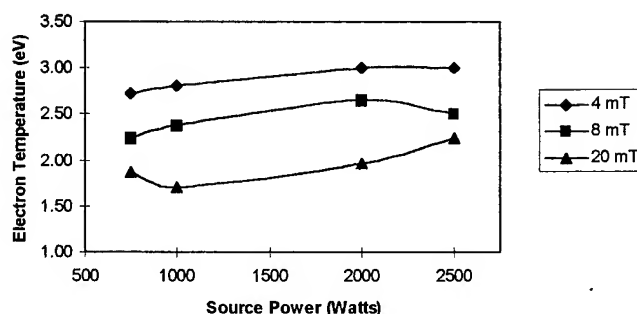


FIG. 3. Electron temperature as a function of source power for pressures of 4, 8, and 20 mT.

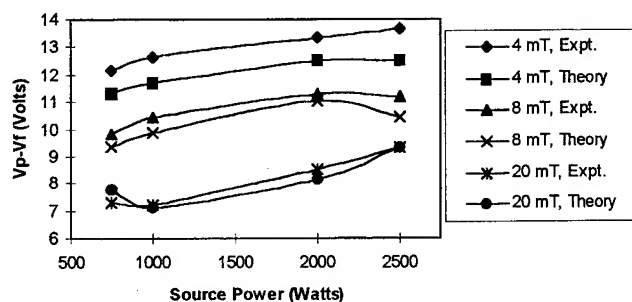


FIG. 4. Experimental and theoretical values of the difference between the plasma and floating potentials as a function of source power for pressures of 4, 8, and 20 mT.

termines the ion energy at the wafer surface) follows the same trend as the electron temperature as one would expect from Eq. (3). There is good agreement between the experimental and theoretical values derived from Eq. (3).

The nitrogen ion density and current density are shown in Figs. 5 and 6. As the source power is raised, more energy is transferred from the Helicon wave to the plasma electrons, the ionization rate increases and the nitrogen ion density increases. As the pressure is raised, the ion density decreases due to a higher ion-electron recombination rate and possibly a lower ionization rate. The ion current density follows the same trend as the ion density as one would expect from Eq. (1). The plasma measurements indicate that one has relatively independent control of the ion energy and ion flux with the pressure and source power, respectively.

V. EXPERIMENTAL NITRIDATION MEASUREMENTS

The goal of the nitridation experiment was to incorporate a high concentration of nitrogen within 1 nm of the silicon dioxide surface in a short period of time. We chose a nitrogen pressure of 4 mT, a source power of 2000 W, and a plasma exposure time of 10 s. The substrate was formed by growing a 4 nm layer of silicon dioxide on a silicon wafer with a furnace process. Following nitridation, depth profiling was performed¹³ by dynamic secondary ion mass spectroscopy (SIMS) using 1 keV Cs primary ion bombardment, and monitoring CsSi^+ , CsO^+ , and CsN^+ ions to track [Si], [O], and [N] concentrations, respectively. Separately, time-of-

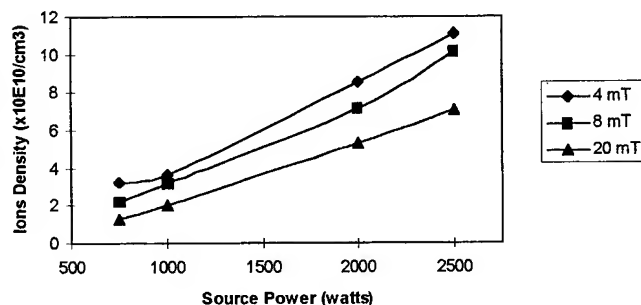


FIG. 5. Nitrogen ion density as a function of source power for pressures of 4, 8, and 20 mT.

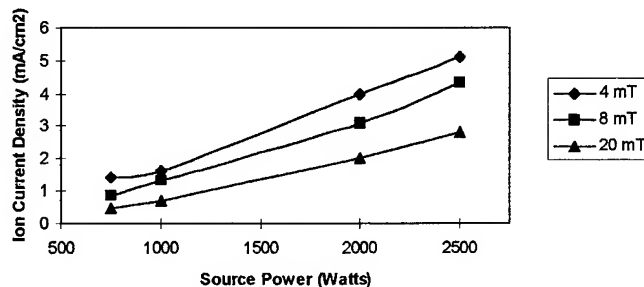


FIG. 6. Nitrogen ion current density as a function of source power for pressures of 4, 8, and 20 mT.

flight SIMS (TOFSIMS) analysis was performed (in house) using 2 keV Ga^+ primary ion bombardment, achieving 0.5–0.7 nm depth resolution within the top 5 nm of the dielectric film.

Figure 7 shows typical depth profiles for nitrogen and oxygen attained with dynamic SIMS. The nitrogen peak occurred approximately 0.5 nm into the silicon dioxide layer and the peak concentration was approximately 17 at. %. TOFSIMS, which has better depth resolution, showed a similar nitrogen and oxygen concentration and range profiles. SIMS measurements from the center and edge of the wafer were nearly identical, reflecting a high degree of plasma uniformity across the wafer as we observed with the Langmuir probe measurements.

An order of magnitude calculation of the number of nitrogen ions incorporated into the silicon dioxide and the number of nitrogen ions which impinged on the silicon dioxide surface indicates that about 1 in 1000 nitrogen ions are actually incorporated into the silicon dioxide. Although this is a small percentage, the very high nitrogen ion flux (2.5×10^{16} ions/s cm^2) results in a peak nitrogen concentration of 17 at. % in only 10 s.

By increasing the source power or the duration of the plasma exposure, the nitrogen concentration was found to increase, although it was not linear with either source power or time. For example, a 60 s plasma exposure resulted in

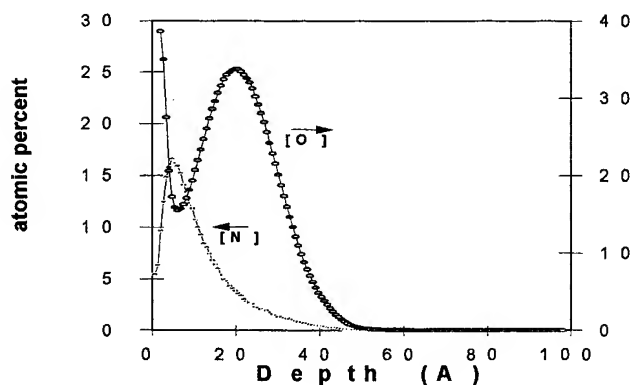


FIG. 7. Atomic nitrogen and oxygen concentrations as a function of position as measured with SIMS.

approximately a twofold increase in the nitrogen concentration. The nitrogen range can be modified slightly by varying ion energy with the pressure, or if desired, a negative bias (rf or pulsed) could be applied to the wafer. We are in the process of further quantifying the dependence of nitrogen dose on exposure time and source power and ion range on pressure.

Transmission electron microscope (TEM) pictures of the nitrided oxide show a well defined nitrided layer at the surface of the oxide. The nitrided layer appears to be amorphous with a well defined interface between the nitrided layer and the silicon dioxide. The TEM pictures showed no visible increase in the surface roughness following the nitridation process as one would expect due to the low ion energy used in the process.

Surface-charge analysis¹⁴ revealed surface voltages of approximately 0.5 V after oxide nitridation which indicates a very low electric field stress on the oxide. The surface voltage was reduced to zero volts after a subsequent postnitridation anneal of the nitrided oxide. Nitrided oxide (100 nm starting oxide, 800 °C N₂O-grown) electrical integrity, as measured by ramped voltage tests on large-area capacitors (0.01 cm²) had a low incidence of low-field breakdowns. Concomitantly, constant current (100 mA/cm²) charge-to-breakdown measurements ranged from 40 to 120 C/cm² for substrate injection, and 0.9–2.4 C/cm² for gate injection, attesting to little damage to the oxide. Details of the electrical measurements can be found in Ref. 3.

VI. CONCLUSIONS

We have demonstrated incorporation of approximately 15 at. % nitrogen into the top 0.5 nm of a silicon dioxide layer with a low pressure, high density nitrogen plasma. The plasma nitridation process incorporates nitrogen much closer to the surface than typical high thermal budget (N₂O or NO) nitridation processes or ion implantation nitridation processes.

Langmuir probe measurements indicate that the nitrogen ion range can be controlled with the process pressure which controls the ion energy. The application of a substrate bias can also be used to further increase the ion range. The nitrogen ion dose can be controlled with the plasma exposure time and the source power which controls the ion flux.

The nitridation process can also be applied to other materials such as silicon to form silicon nitride. Silicon nitride formed in this manner would have a low hydrogen concentration (unlike conventionally formed silicon nitride) which would make it a gate dielectric candidate for CMOS transistors.

ACKNOWLEDGMENTS

The authors would like to thank A. Templeton (MSL) for his assistance with the TOFSIMS analysis and D. Springfield for assistance with the Langmuir probe setup.

¹S. V. Hattangady, H. Niimi, and G. Lucovsky, *Appl. Phys. Lett.* **66**, 19 (1995).

²S. R. Kaluri and D. W. Hess, *Appl. Phys. Lett.* **69**, 19 (1996).

³S. V. Hattangady, R. Kraft, D. T. Grider, M. A. Douglas, G. A. Brown, P. A. Tiner, J. W. Kuehnc, P. E. Nicollian, and M. F. Pas, *IEDM Tech. Dig.*, p. 495 (1996).

⁴D. T. Grider, S. V. Hattangady, R. Kraft, P. E. Nicollian, J. Kuehne, G. Brown, S. Aur, R. H. Eklund, M. F. Pas, W. R. Hunter, and M. Douglas, *VLSI Tech. Dig.*, p. 47 (1997).

⁵P. J. Tobin et al., *J. Appl. Phys.* **75**, 1811 (1994).

⁶Z.-Q. Yao, *Appl. Phys. Lett.* **64**, 3548 (1994).

⁷N. Hershkovitz, in *Plasma Diagnostics, Vol. I, Discharge Parameters and Chemistry*, edited by O. Auciello and D. L. Flamm (Academic, Berkeley, CA, 1989), pp 113–183.

⁸PINNACLE Etch System, Plasma and Materials Technology Inc., Chatsworth, CA 91311.

⁹R. W. Boswell, *Phys. Lett. A* **33**, 457 (1970).

¹⁰F. F. Chen, *J. Vac. Sci. Technol. A* **10**, 1389 (1992).

¹¹Smart Probe Scientific Systems, Ltd., DCU, Glasnevin, Dublin 9, Ireland.

¹²F. F. Chen, in *Plasma Diagnostic Techniques*, edited by R. H. Huddleston and S. L. Leonard (Academic, New York, 1965), pp. 113–199.

¹³SIMS measurements performed by Evans East, Plainsboro, NJ 08520.

¹⁴A. M. Hoff, *Solid State Technol.* **39**, 139 (1996).

Ar, N₂, and Cl₂ electron cyclotron resonance plasmas measured by time-of-flight analysis: Neutral kinetic energies and source gas cracking

Rory S. Goodman, N. Materer, and Stephen R. Leone^{a)}

JILA, National Institute of Standards and Technology, University of Colorado, and Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309

(Received 23 February 1997; accepted 30 May 1997)

Neutral mean kinetic energies, ion intensities, and neutral source gas cracking from Ar, N₂, and Cl₂ electron cyclotron resonance (ECR) plasmas, are measured by modulated beam time-of-flight (TOF) analysis. The TOF distributions are characterized by a two component form consisting of an effusive Maxwell-Boltzmann distribution and a fast Gaussian component, that accounts for nonthermal species produced in the source. The mean kinetic energies of neutral species are found to range between 0.04 and 0.45 eV, depending on species and plasma conditions. Mean kinetic energies increase at a nearly constant rate, with decreasing pressure from 8.0×10^{-2} to 2.5×10^{-2} Pa with constant applied microwave power. At pressures below 2.5×10^{-2} Pa, the neutral mean kinetic energies sharply increase. This sharp increase in neutral mean kinetic energy is attributed to an abrupt increase in the ion flux out of the source. The increase in kinetic energy can be separated into two contributions, (i) thermal at higher pressures and (ii) nonthermal at lower pressures. This effect is much stronger for atomic neutrals than for molecular neutrals, where internal degrees of freedom can accept energy in momentum transfer collisions. Cracking of N₂ and Cl₂ is also examined as a function of source pressure at constant microwave power. The N:N₂ flux ratio from the ECR source varies between 0.2 and 1.4. The Cl:Cl₂ flux ratio varies from 10 to 16, indicating a very high degree of dissociation in the plasma. Both flux ratios decrease with increasing source pressure. The total flux of Cl increases with increasing source pressure over the entire range, while the N atom flux peaks at $\sim 1.3 \times 10^{-2}$ Pa and decreases on either side of this pressure.

© 1997 American Vacuum Society. [S0734-211X(97)14904-6]

I. INTRODUCTION

The production of ultralarge scale integration (ULSI) circuits entails precise control over morphological features transferred to various materials on a submicron scale. The use of plasmas, which provide energetic and chemically active species to etch micron and submicron features, is ubiquitous in the production of solid state devices.¹⁻⁵ As circuit element densities are increased to improve device performance, higher degrees of etch anisotropy and selectivity are required. Electron cyclotron resonance (ECR) plasma sources, with their high reactive particle and ion densities, have been suggested as a potential means of achieving greater control over device morphologies, in many etching and deposition applications.¹

ECR plasmas consist of a coupled electron, ion, and neutral gas plasma with superimposed static magnetic and microwave fields. Energy is resonantly absorbed by free electrons from the applied microwave field when their cyclotron frequency, which is determined by the local magnetic field strength, is equal to the frequency of the microwave field. Typically, microwaves at 2.45 GHz are employed, which results in ECR coupling at a magnetic field strength of 875 G. These ECR heated electrons then transfer energy to heavy particles through collisions and through the action of ambipolar drift fields. This ECR coupling allows efficient power absorption at low pressure, which results in a plasma with

high radical and ionic particle densities.¹ These reactive and energetic particles are responsible for the etch and deposition characteristics, which include etch rates, etch anisotropy, and film morphology.⁶⁻⁸ A detailed understanding of the behavior of ECR plasmas as a function of process parameters, may lead to greater control over etch and deposition characteristics.

The many potential applications of ECR plasma sources, have motivated considerable experimental effort aimed at determining how excited particles produced in ECR plasmas depend on external parameters. Kinetic energy distributions of ionic and neutral species in ECR plasmas have been measured by laser induced fluorescence (LIF) Doppler spectroscopy,⁹⁻¹⁴ emission Doppler spectroscopy,^{9,15-18} absorption Doppler spectroscopy,¹⁹ and nonoptical methods.²⁰⁻²² Emission Doppler linewidth measurements have obtained neutral and ion average kinetic energies ranging between 0.1–0.6 and 1.0–2.5 eV, respectively.¹⁸ However, other groups report neutral and ion temperatures of only a few tenths of an eV, with neutral temperatures generally cooler than ion temperatures having energies as low as 0.03 eV.⁹⁻¹⁷ These results are summarized in Table I. The considerable spread in reported values may be due to the different plasma reactor geometries and conditions, since ECR plasma properties depend on various parameters in a complex interdependent fashion. Additionally, Doppler measurements are difficult to perform and require a detailed knowledge of the instrument response function and other possible contributions to the linewidth in order to extract

^{a)}Staff member, Quantum Physics Division, National Institute of Standards and Technology; Electronic mail: SRL@JILA.colorado.edu

TABLE I. Summary of previous kinetic energy measurements of species in ECR plasmas.

Plasma species	Average translational energy/eV	Experimental technique	Reference
Ar	0.041–0.072	LIF, emission Doppler	9
Ar	~0.5	Emission Doppler	16
Ar ⁺	<0.1		
Ar	0.1–0.6	Emission Doppler	18
Ar ⁺	1.0–2.5		
Ar	0.068, ^a 0.030 ^b	LIF Doppler	12
Ar ⁺	≤0.5, ^c ~0.25 ^b		
Ar,He	0.032±0.004 to 0.072±0.006	Pressure depletion	13
Ar ⁺	0.26–0.47	LIF Doppler	
Ar	0.2	Pressure depletion	20
Ar	10–20	TOF	21
N ₂ ⁺	0.12–0.25	LIF Doppler	11
Cl	0.069±0.009	Absorption Doppler	19
Cl ⁺	0.24–0.61, ^d 0.26–0.42 ^c	LIF Doppler	14

^aIn the ECR plasma.^bDownstream from the ECR plasma.^cParallel to ECR magnetic field.^dPerpendicular to the ECR magnetic field.

accurate velocity information. Finally, some measurements are made directly within the plasma, while others interrogate the species emanating from the plasma.

A previous study by Tsuchizawa *et al.* employed modulated beam time-of-flight (TOF) mass spectrometry, to measure the flux of neutral Ar emanating from a pin hole downstream from an Ar ECR plasma.²¹ This group observed some fast neutrals with median energies between 10 and 20 eV, which is approximately the energy of the ions emanating from the plasma, suggesting that some fast neutrals are produced directly from ions via a charge exchange process.²¹ However, in this study the ion extraction voltages of the quadrupole mass spectrometer, were adjusted to discriminate against slow particles in order to reduce the Ar background contribution to the mass signal. Therefore, the efficiency of the fast neutral production and the fraction of fast neutrals, compared to all neutrals emanating from the source were not determined.

Theoretical methods have also been employed to obtain an understanding of ion and neutral transport in high-density plasmas, i.e., plasmas that have high ion and radical densities.^{23,24} Monte Carlo (MC) simulations of low pressure, high ion density plasmas suggest that hot neutrals produced through charge exchange, and electron impact dissociation will be present at wafer surfaces with fluxes comparable to those of ions.²⁴ These hot neutrals could contribute to activated surface processes, that otherwise would not be accessible, such as isotropic etching.²⁴ Other MC simulations suggest significant neutral density depletion in and downstream from an ECR plasma source, due to neutral ionization and collisional heating of neutrals.²³

The degree of source gas dissociation and the flux of dissociated particles at the wafer surface are also of central importance to the etching and deposition characteristics of

ECR plasmas. However, few studies have quantified these details for even the most simple systems, such as N₂ and Cl₂. Chlorine dissociation in an ECR plasma has been examined through both kinetic simulation, emission, and LIF spectroscopies.^{25,26} They report the Cl/Cl₂ ratio in their ECR plasma is 1%–2%, suggesting little cracking of Cl₂ source gas.²⁵ Additionally, their results from simulations suggest that the Cl₂⁺ molecular ion is more abundant than the atomic ion, Cl⁺, in ECR plasmas.²⁶ In contrast, another study modeling non-ECR radio frequency and microwave chlorine discharges, suggests a high degree of chlorine dissociation.²⁷ Dissociation of nitrogen in an ECR plasma has also been examined through emission spectroscopy and chemical titration.^{28,29} The chemical titration study measured a N/N₂ dissociation ratio of 12% downstream from an ECR plasma.²⁸ The emission study, performed in the same type of ECR source as this study, observed strong emission from N atoms, suggesting significant cracking of N₂, but the dissociation fraction was not quantified.²⁹

This article describes the use of modulated beam TOF analysis to measure neutral mean kinetic energies, ion production, and source gas cracking of species emanating from Ar, N₂, and Cl₂ ECR plasmas. These TOF measurements reflect the velocity components, perpendicular to the wafer surface rather than the velocities of species in the bulk plasma. In all the experiments performed in this study, the measurements are made without any applied accelerating potentials to extract ions from the ECR source. Many studies of Ar ECR plasmas have been performed due to the relative simplicity of this system, therefore, measurements of Ar ECR plasmas are performed as a benchmark for comparison with the earlier studies. N₂ and Cl₂ plasmas are chosen for study because of their relevance to plasma-enhanced deposition of III–IV materials and etching, respectively. The relationship between source pressure and neutral mean kinetic energy for Ar, N₂, and Cl₂ ECR plasmas at constant microwave power is reliably obtained. Additionally, a strong correlation between increased ion flux and increased kinetic energy of neutral particles, at lower pressures is observed. Cracking of N₂ and Cl₂ source gases and relative atom fluxes as a function of source pressure are also reported.

II. EXPERIMENT

Modulated beam TOF analysis is employed to study neutral and ionic species that emanate from Ar, N₂, and Cl₂ ECR plasmas. A continuous beam of energetic particles generated in the ECR plasma source, passes from this source to a slotted disk chopper, where the beam is modulated. The particles that make up the modulated beam, then pass from the chopper to a quadrupole mass spectrometer (QMS), where they are detected. The elapsed flight times of the particles in the modulated beam between the chopper and the QMS are recorded, resulting in a TOF distribution. This TOF distribution is used to determine the mean kinetic energy and intensity of neutral species downstream from the ECR plasma. Ion intensities are also measured by this method and exhibit correlated behavior with neutral measurements.

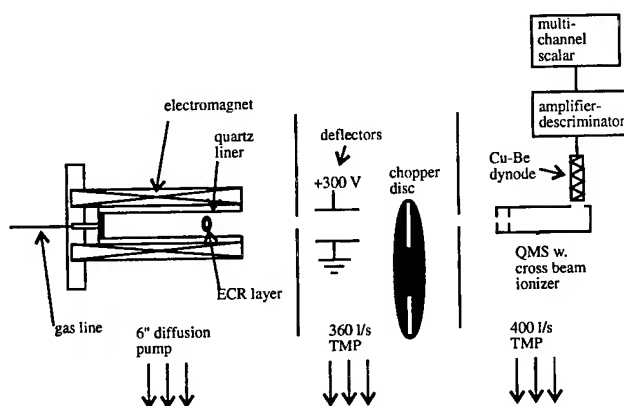


FIG. 1. Schematic of modulated-beam TOF-QMS system illustrating the ECR source, chopper and QMS chambers.

A schematic of the experimental apparatus used in this study is shown in Fig. 1. The apparatus consists of three chambers, one each for the ECR source, chopper, and QMS. The energetic particles of interest are generated in an ECR (AsTex compactTM ECR plasma source), plasma source powered by a 250 W 2.45 GHz microwave power supply. The microwave power supply indicates total microwave power output and the reflected power. A slug tuner allows efficient coupling of microwaves, and is adjusted to minimize the reflected microwave power to <10 W. The spatially varying static magnetic field is produced by a water-cooled solenoid magnet, surrounding the plasma chamber. This electromagnet is powered by a 40 V, 25 A dc power supply. This arrangement results in a configuration in which the magnetic field strength decreases, continuously from the back of the source to the source exit. The ECR region, with a magnetic field strength of 875 G, may be shifted relative to the source exit by varying the magnet current. The plasma source was operated with an applied microwave power (forward power minus the reflected power) of 200 W, and a magnet current of 17 A in all of the experiments reported here. These conditions result in efficient microwave coupling to the plasma and placed the ECR layer 3.8 cm from the source exit. Source gases are introduced through the rear of the source around a cylindrical ceramic plug, which also acts as the waveguide for launching microwaves into the plasma chamber. The plasma chamber consists of a quartz liner ~ 2.5 cm in diameter and 15.2-cm-long. The source chamber pressure during operation varies from 6.7×10^{-3} to 9.3×10^{-2} Pa, depending on the source gas flow rate. Conductance calculations reveal that the pressure in the source, close to the ECR layer, is approximately an order of magnitude higher than that of the surrounding chamber. Source gas flow rates are monitored using flow meters. Ar, N₂, and Cl₂ source gases are 99.97% purity or greater and are used without further purification.

The chopper chamber is separated from the source chamber by a grounded stainless steel plate with a 2.0 mm aperture. This chamber contains a slotted chopper disk driven by a 61 Hz electric motor. The chopper disk has two sets of slits

180° apart. Each set of slits has three sections with different widths corresponding to 50, 250, and 550 μ s openings at a disc rotation of 61 Hz. The widest slit is required for all measurements in order to obtain sufficient signal-to-noise. The chopper assembly is mounted on a translation stage; so that the chopper disc may be moved transversely to the source aperture, allowing the particles leaving the source to pass through any of the three slit widths. The edges of each slit are cut along the radius to give a constant width in time over the length of each section. The aluminum chopper disk is grounded with a copper brush, that makes contact to the motor axle on which the chopper is mounted. Also contained in this chamber is a set of deflectors used to remove ions from the beam, when neutrals alone are of interest.

The last chamber is separated from the chopper chamber by a grounded 2.5 mm stainless steel aperture. A QMS, with an off axis electron multiplier and a crossed beam ionizer, is mounted in this last chamber and is used to detect the modulated atoms and molecules from the ECR plasma source. The electron energy in the crossed beam electron bombardment ionizer is set to 70 eV. The flight length from the chopper to the ionizer of the QMS is 0.41 m. This is followed by a second flight path through the QMS to the detector of 0.25 m. Ions are extracted from the cross beam ionizer with an 11.8 V potential between the ionization region and the rods of the QMS. The base pressure of the QMS chamber is 1×10^{-6} Pa and remains $<1 \times 10^{-5}$ Pa, while the ECR source is in operation. Even with a base pressure of 1×10^{-5} Pa, the background level of the species of interest was relatively high and required the co-addition of many chopper cycles to acquire a TOF distribution with sufficient signal-to-noise.

A beam from the source is defined by the aperture between the source and chopper chambers, and the entrance aperture of the ionizer of the QMS. The apertures are aligned before the experiments using a theodolite. The QMS is operated in a pulse counting mode and the signal is collected by a multichannel scalar, which is triggered by a light-emitting diode (LED) photodiode switch mounted on the slotted disk chopper assembly. The LED-photodiode switch is mounted such that the multi-channel scalar is triggered prior to the chopper slit's pass, through the beam to ensure the entire transmission of gas is measured. The multichannel scalar has a record length of 1024 bins, each with a width of 6 μ s to give a total collection time of 6.144 ms. This is sufficient to collect the entire TOF distribution and allows the baseline to flatten before the end of the record. TOF distributions are collected via 50 000 co-additions at 122 Hz; each co-addition corresponds to one opening of the chopper through the beam. Data are saved to disk for subsequent analysis. Data analyses of TOF distributions will follow in Sec. III. Slotted disk velocity measurements have been described in detail by many authors and a detailed discussion of the technique is available.³⁰

III. RESULTS AND DISCUSSION

The following sections describe the acquisition and analysis of TOF data for atomic and molecular species emanating

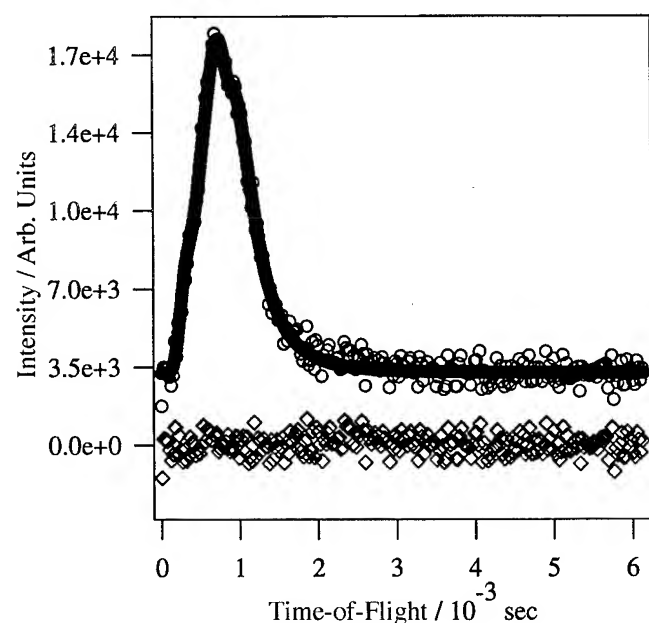


FIG. 2. Typical neutral Ar TOF distribution from the ECR source after subtraction of the component, due to UV emissions and the background of the detector chamber. A fit to the data and the residuals are also pictured.

from Ar, N₂, and Cl₂ ECR plasmas. Acquisition of TOF distributions, determination of neutral mean kinetic energies, ion intensities, and treatment of cracking data for molecular source gas plasmas are discussed in detail. A discussion of results is included in each section.

A. Argon

Figure 2 shows typical TOF data for neutral Ar from an Ar ECR plasma, along with a fit to the data (described below), and the residuals from this fit. This spectrum is obtained by a subtraction of two scans of 50 000 co-additions each. Typical raw co-addition traces from which TOF distributions are extracted and are depicted in Fig. 3. The QMS was set to detect Ar⁺ with $m/e=40$. The first scan, Fig. 3(a), is taken with the ionizer of the QMS on and the deflectors in the chopper chamber biased to +300 V, to remove ions from the beam. This scan includes contributions from ultraviolet (UV) source emissions, Ar gas background in the QMS chamber and directed neutrals born in the ECR plasma. The second scan, Fig. 3(b), is taken with the deflectors biased and with the QMS ionizer off. The signal in this scan is due mainly to UV emissions from the ECR plasma, which generate secondary electrons in the QMS. The Ar TOF data in Fig. 2 is generated from a subtraction of the two scans, Fig. 3(a) minus Fig. 3(b), followed by a background subtraction. The background level, which is due to the gas background in the QMS chamber, is fit and subtracted during the fitting procedure described below. The last trace in Fig. 3(c) shows the result, when the deflectors are grounded and the ionizer is off. This condition passes the ions and the UV light from the source. The difference of Fig. 3(c) minus Fig. 3(b) is the ion flux.

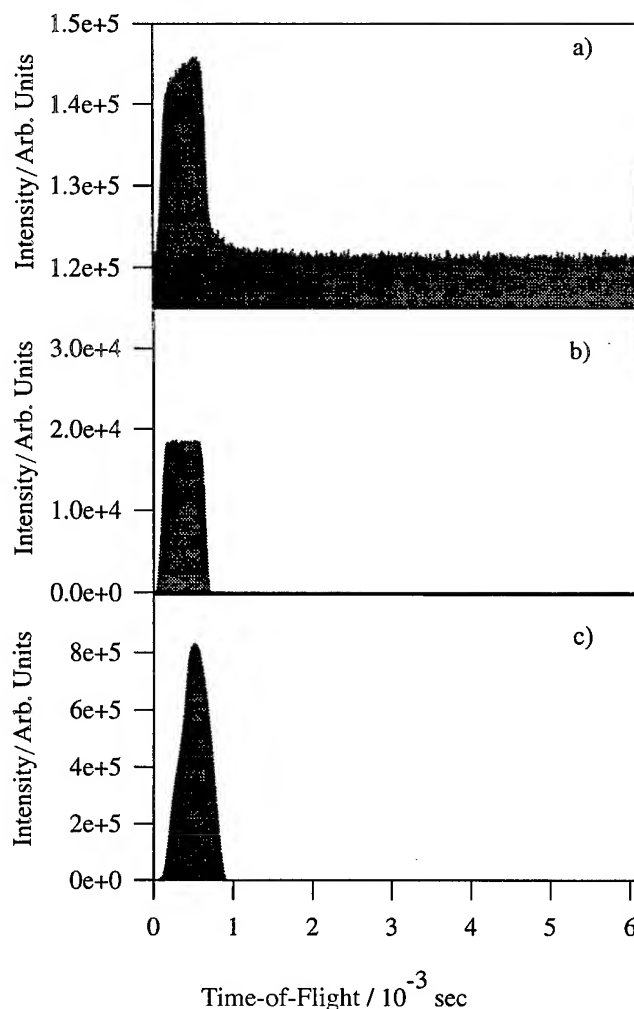


FIG. 3. Raw data used to acquire the TOF distribution and ion flux intensity for Ar from the ECR source with the source pressure below 1.5×10^{-2} Pa. 50 000 passes of chopper with, (a) deflectors biased to 300 V and the ionizer of the QMS on; (b) deflectors biased and ionizer off; (c) deflectors grounded and ionizer off. The ordinate axis in (a) extends to zero, but is not shown for clarity. The neutral TOF distributions were determined by subtracting (b) from (a), a typical result of which is seen in Fig. 2. The ion signals are determined by subtracting (b) from (c) and integrating the resulting difference. The UV component in panel (b) represents a worst case, and is much smaller at higher pressures and with other source gases.

The resulting neutral TOF distribution, $G(t)$, is not a pure reflection of the velocities of neutrals leaving the source, but is a convolution of the true TOF, $g(t)$, with the chopper slit function, $ch(t)$:³⁰

$$G(t) = \int g(t-x)ch(x)dx. \quad (1)$$

A close approximation to the chopper function is shown in Fig. 3(b), and is obtained from many co-additions of the UV emission, when the ionizer is left off and the deflectors are biased to exclude ions. As the slit of the chopper disc passes through the solid angle of collection of the QMS, defined by the first aperture in front of the plasma source, and the aperture of the ion source on the QMS, the signal from UV emissions gives a close approximation to the shape

of the chopper function. In principle, a slit with a very narrow width compared to the TOF distribution could be used to obtain a measurement close to the actual TOF distribution. However, the signal is sufficiently weak, due to a high background of species in the mass spectrometer detection chamber, that the widest slit, with an opening time of 550 μ s, is required for all neutral measurements. As a result, the chopper width is comparable to the width of the TOF distributions; therefore, the precise shape of the true TOF distribution, $g(t)$, is lost, and is represented by $G(t)$, in Eq. (1).

Mean kinetic energies of atomic and molecular species are calculated from mean TOF, determined from the TOF distributions. The mean time $\langle t \rangle$ is defined by

$$\langle t \rangle = \frac{\int_0^\infty t G(t) dt}{\int_0^\infty G(t) dt}, \quad (2)$$

where $G(t)$ is the TOF distribution defined above. Determining $\langle t \rangle$, directly from the data proves to be difficult, since any noise in the data makes $\langle t \rangle$ very sensitive to the position of the baseline of $G(t)$. With a perfect choice of base line, the random scatter of data around the base line, would not skew the determination of the mean; however, if the base line were slightly over or underestimated, the numerator in Eq. (2), and hence the mean value, $\langle t \rangle$, would be incorrect. In order to avoid this problem, the mean value and integrated neutral intensity are determined from a fit to the data. An attempt to fit the measured TOF distribution with an effusive Maxwell-Boltzmann distribution, with and without a stream velocity and convoluted with the chopper function, was performed and resulted in inadequate agreement with the data. Therefore, a two component form shown in Eq. (3), consisting of a Gaussian at short times and an effusive Maxwell-Boltzmann distribution without a stream velocity is used (see Ref. 30 for a derivation of this component):

$$g(t) = a \exp\left(-\frac{(t-t_0)^2}{\delta}\right) + b \frac{1}{t^4} \exp\left(-\frac{ml^2}{2kTt^2}\right) + c. \quad (3)$$

Equation (3) contains three constants: k , the Boltzmann constant, m , the mass of the particles being measured, and l , the flight length between the chopper and entrance to the QMS. This form is convoluted with the chopper function, Fig. 3(b), and fit to the data by minimizing χ^2 with a Powell minimization routine as a function of the six remaining parameters.³¹ These parameters are a scale factor for the Gaussian and Maxwell-Boltzmann components, a and b , an offset, c , a temperature, T , and the Gaussian center, t_0 , and width, δ . An example of a fit generated by this procedure is shown as the solid line in Fig. 2. Due to the width of the chopper opening and the large number of fit parameters, no attempt was made to draw any physical information from the fit parameters. As mentioned above, the purpose of the fit is to aid in computing mean kinetic energies, and integrated intensities of particles in the beam from the shape of the distribution without complications due to noise in the data.

Using these fits to the TOF data, mean flight times for particles emanating from the source are calculated. In order to obtain an accurate mean TOF, the flight time through the

QMS must be accounted for, and the time base must be adjusted so that $t=0$ corresponds to the middle of the chopper opening. In order to do this in a simple fashion, an approximation is made that the width of the TOF distribution does not increase during the flight through the QMS. In a more rigorous treatment, the initial velocities of particles entering the QMS should be accounted for, since they will cause the distribution to spread in time during flight through the QMS. However, this effect does not need to be considered here, because the initial velocity of most particles entering the QMS is small compared to the velocity after extraction from the ion source. Ions formed by electron impact ionization in the ionization region are extracted with an 11.8 V potential, between the ionizer and the rods of the QMS. The difference in flight time through the QMS for an Ar⁺ ion with an initial translational energy of 0.03 eV (thermal) and an Ar⁺ ion with an initial translation energy of 20 eV, which is estimated to be an upper limit on the energy of fast particle exiting the ECR source, is only 12.7 μ s, which is just greater than the width of two bins of the multichannel scalar. The arrival distributions are generally 1800- μ s-wide; hence, the width of the TOF distributions will be almost unaffected by the flight through the QMS. Under this approximation, the mean TOF is first shifted by the time it takes for an ion accelerated by the 11.8 V extraction voltage, to pass through the 0.25 m path through the QMS. For an Ar⁺ ion with no initial velocity this flight time in the QMS is 33.2 μ s. Next, the mean TOF is shifted by the amount necessary to adjust the $t=0$ to the middle of the chopper opening, which is exactly 348 μ s from the actual trigger point. The mean TOF, calculated from the measured TOF distribution, shifted by the two corrections mentioned above, closely approximates the mean TOF of particles over the 0.41 m between the chopper and the ionizer of the QMS. These mean flight times are then converted into mean kinetic energy.

To verify the validity of this procedure, measurements were made on thermal argon beams generated by backfilling the source chamber with argon. The gases are introduced via a tube off the beam axis. A thermal beam defined by the aperture between the source and the chopper chambers was modulated by the chopper and measured by the technique described above. Mean TOF measurements result in mean kinetic energies of 0.026 ± 0.001 eV. These experiments are also conducted by introducing gas through the source, where the inlet orifice lies on the beam axis. These measurements resulted in mean translational energies of 0.033 ± 0.001 eV, indicating flow from the orifice in the source is not purely effusive. This noneffusive behavior manifests itself in the measurements of neutral intensities, and is further discussed below for nitrogen and chlorine measurements. All data presented below are measured for gases introduced through the source orifice.

Figures 4(a) and 4(b) show the mean kinetic energies of neutral Ar emanating from Ar ECR plasmas with slightly different conditions, as noted below. The mean kinetic energies are determined as indicated above. The coincident Ar⁺ ion intensity emanating from the ECR plasma source

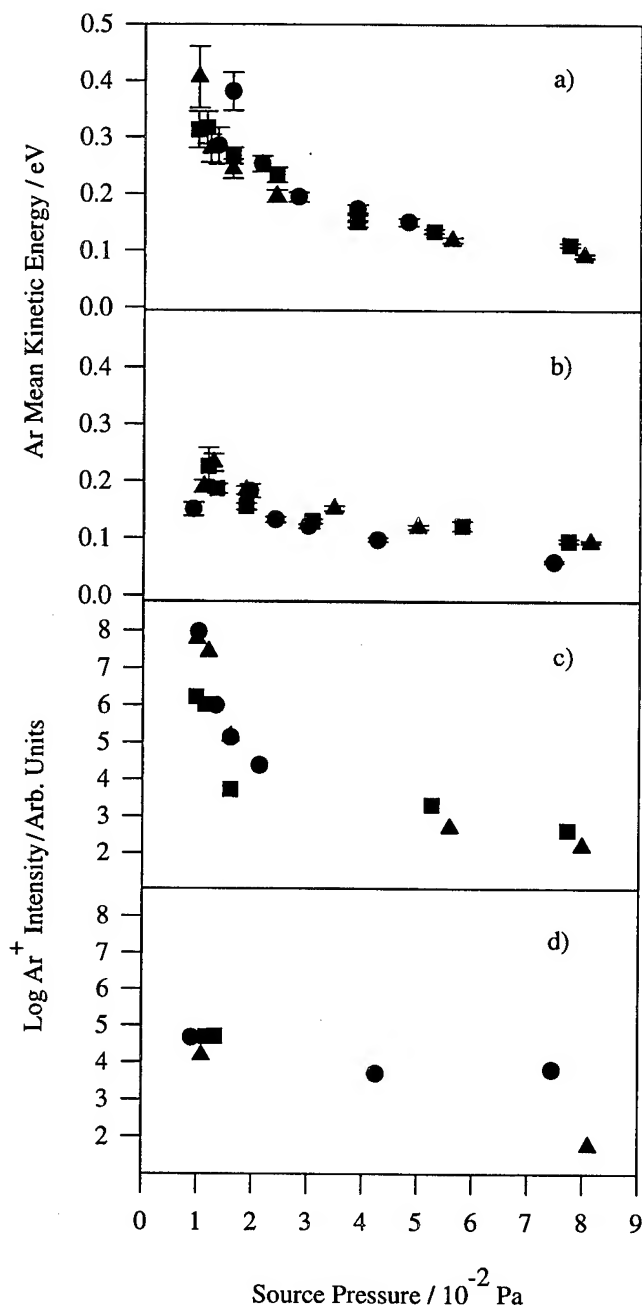


FIG. 4. Mean neutral kinetic energy of Ar, (a) and (b), and ion flux intensity of Ar⁺, (c) and (d) from Ar ECR plasmas vs source chamber pressure. ●, ■, ▲ indicate data taken on different days. Panels (a) and (c) are measured from the same plasma. Panels (b) and (d) are from another plasma, where the impedance matching was purposely detuned to reduce ion flux. Applied microwave power of 200 W and magnet current of 17 A for both sets.

versus source chamber pressure, measured for the same two plasmas as in Figs. 4(a) and 4(b), are shown in Figs. 4(c) and 4(d), respectively. The ion intensities were determined by integrating the ion flux, the determination of which is described above. However, ion energies were not calculated from the TOF distributions, since the mechanical chopper and the voltages in the QMS significantly alter their flight times. The ion energies would be better measured by a retarding field analyzer and are left to future work. As

mentioned above, the applied microwave power and magnet current are held at 200 W and 17 A, respectively. For the measurements shown in Figs. 4(b) and 4(d), the slug tuner, used to impedance match the plasma to the microwave source, is slightly detuned. This increased the reflected power; however, the absorbed microwave power is maintained at 200 W by increasing the forward power. As indicated above, the pressure in the source around the ECR layer is estimated to be about an order of magnitude higher than the source chamber pressure indicated on these plots. Data for mean kinetic energies and ion intensities are taken together, before changing plasma parameters, ensuring that plasma conditions were as identical as possible for the neutral and ion measurements.

The Ar neutral kinetic energy continuously increases at low pressures below $\sim 1.5 \times 10^{-2}$ Pa only when accompanied with a sharp increase in the ion flux, as indicated in Fig. 4. Above 1.5×10^{-2} Pa the ion signal is very low and the increase in neutral kinetic energy is much more gradual. The two dominant processes determining the neutral mean kinetic energy are charge exchange between fast ions and the neutral background and the moderation of fast ions and neutrals through elastic collisions with the slower neutral background. At high pressures, the plasma is close to thermal because most particles undergo many collisions before exiting the source. One would expect, by an energy balance argument, that the mean kinetic energy in this higher pressure regime would increase as the pressure and, hence, the density is lowered at constant absorbed microwave power; this is observed. The plasma is thought to be under this near-thermal condition at source chamber pressures above 1.5×10^{-2} Pa. As the pressure is decreased below 1.5×10^{-2} Pa, the mean free path in the source becomes long enough that the particles in the plasma do not approach equilibrium before exiting the source region. At the same time, the measured flux of ions out of the source, shown in Fig. 4(c), greatly increases as the pressure is lowered. In this low pressure regime, the number of fast neutrals that are born via charge exchange increases, raising the mean neutral kinetic energy. Direct evidence supporting charge exchange as a mechanism for neutral heating was provided by Tsuchizawa *et al.*, where neutrals with kinetic energies of 10–20 eV were measured.²¹ We found evidence for such high velocity neutrals at the leading edge of the TOF distributions measured in this study; however, it is clear from mean kinetic energy data presented here that most of the neutrals emanating from the source are much less energetic. Additionally, the cross section for resonant charge transfer with collision energies between 4 and 25 eV is the same order of magnitude as the elastic cross section.³² Thus, many fast neutrals generated through charge exchange will also undergo elastic collisions with the Ar background, redistributing energy among the neutral background.

The measurements shown in Figs. 4(b) and 4(d) are analogous to those shown in Figs. 4(a) and 4(c), but, as mentioned above, with different plasma conditions. In contrast to the data in Figs. 4(a) and 4(c), the neutral mean kinetic energy of

Fig. 4(b) does not rise as sharply at low pressure, nor does the corresponding ion signal in Fig. 4(c), supporting the correlation between higher neutral mean kinetic energies and the flux of ions out of the plasma. The origin of the slug tuner's effect is not understood. The microwaves could be absorbed by some mechanism other than ECR heating, reducing the effective power absorption and, hence, the ion density. The spatial variation of the microwave absorption could also change, resulting in lower ion flux at the sampling aperture. However, the reduction in ion flux, whatever the root cause, is clearly correlated with the neutral mean kinetic energies.

Table I summarizes the previous measurements of average kinetic energies of various plasma species produced in a variety of ECR plasma systems. The results for neutral Ar average kinetic energies vary over a considerable range, but generally begin at 0.03 eV and range up to 0.6 eV. This study reports, neutral kinetic energies from ~ 0.1 to 0.41 eV for neutral Ar emanating from an Ar ECR plasma. Optical emission results from two groups show that the neutral gas temperature decreases with decreasing source pressure.^{15,18} One difference between these measurements and those of this study, is that the neutrals pass through a sheath formed on the sampling aperture in front of the plasma, whereas those measured optically are in the bulk plasma and do not cross a sheath region. The measured neutral mean kinetic energy is sensitive to the ion flux through the aperture bordering the plasma, which can provide additional neutral heating through charge exchange collisions. Additionally, the previous studies examined plasmas at higher pressure and with different reactor geometries.

B. Nitrogen

1. Mean neutral kinetic energies

Figures 5(a) and 5(b) show the N and N₂ neutral mean kinetic energies versus source chamber pressure with an applied microwave power and magnet current of 200 W and 17 A for a nitrogen ECR plasma. These measurements were made in an identical manner to the Ar neutral mean kinetic energy measurements detailed above. However, there is a contribution in the atomic TOFs from atoms that were produced via dissociation of the parent molecule in the ionizer of the QMS rather than in the plasma. This contribution will result in a slight underestimation of the N atom mean kinetic energies. This error is small since the cracking in the QMS is small compared to cracking in the ECR plasma as indicated by the cracking data below. The thermal mean kinetic energy data are also shown in Figs. 5(a) and 5(b) and are indicated by open symbols. This thermal data were analyzed in the same manner as that for Ar. Figure 5(c) shows relative N⁺ and N₂⁺ ion intensities versus source chamber pressure measured at the same time as the neutral data shown in Figs. 5(a) and 5(b). The ion intensity of N⁺ has been reduced by a relative quadrupole mass transmission efficiency, $s_N/s_{N_2} = 1.2 \pm 0.1$, the determination of which is discussed in Sec. III B 2.

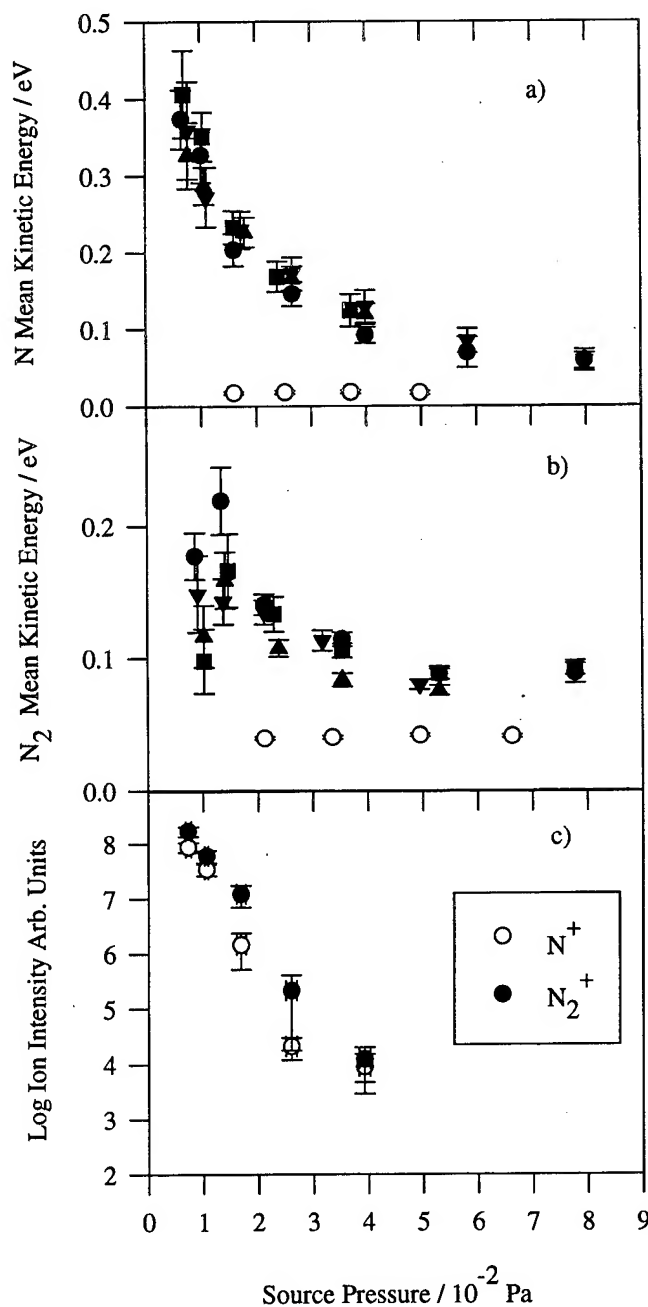


Fig. 5. Neutral mean kinetic energy for N, (a), and N₂, (b) vs source chamber pressure. ●, ■, ▲, and ▼ indicate data taken on different days with the plasma on. ○ indicate data with the plasma off. Ion flux intensity, (c), vs source chamber pressure for the same plasmas. Ion data from different days was averaged for clarity. Microwave power of 200 W and magnet current of 17 A.

The mechanism involved in the production of fast neutrals in the nitrogen plasma is similar to that in the Ar plasma discussed above. However, dissociation of molecular N₂ to form N atoms provides an additional route for the production of fast neutral N.³³ The other possible mechanism for the fast neutral production is through charge exchange collisions between N⁺ and N atom. The data suggest that charge exchange may be important at the lower pressures, since the increase in ion intensity in Fig. 5(c) is accompanied by an

increase in N mean kinetic energy in Fig. 5(a). One would expect that the detection of fast neutrals produced through charge exchange reactions would be more likely in this type of measurement, since they will be directed by the ion velocity, which has a large component perpendicular to the sampling aperture due to the ambipolar fields along the magnetic field lines produced by the source. In contrast, the energetic neutrals produced through dissociation will be directionally isotropic, reducing the efficiency with which they will emerge from the source chamber relative to those produced through charge exchange with directed ions.

The mean kinetic energy of the N₂, shown in Fig. 5(b), emanating from the source is similar to that of the N mean energies; however, the kinetic energies at low pressure do not rise as rapidly for N₂ as for N atom. At the higher pressures investigated, the mean kinetic energies of both N and N₂ are about equal, suggesting that they may have reached equilibrium before exiting the source, however, at lower pressure this is not the case. To our knowledge, the kinetic energies of neutral N and N₂ produced in ECR plasmas have not been previously measured; however, a mean kinetic energy of 0.12–0.25 eV for N₂⁺ produced in an ECR plasma has been measured by Den Hartog *et al.*¹¹ These results seem consistent with ion energies in the bulk plasma measured for other species summarized in Table I.

2. Source gas cracking

In addition to mean kinetic energies, cracking of the molecular source gas in the ECR source is also investigated. Figure 6(a) shows the cracking ratios of neutral nitrogen, as a function of source chamber pressure determined from the measured N and N₂ signals with the plasma on and off. The intensities of N and N₂ were determined by integrating the area under the fits of the neutral TOF data. No attempt has been made to include ion intensity data with the neutral data in the cracking results, since the neutral and ion data will not be attenuated equally as the particles travel from the aperture in front of the plasma to the QMS. Ions accelerated across the sheath formed at the sampling aperture in front of the source will be more directed than most of the neutrals, which will be more effusive in nature. However, the ions will also have losses due to Coulombic repulsion between ions in the beam and stray fields that may exist along the flight path. Hence, the intensities of ions and neutrals measured by the QMS do not directly reflect the relative fluxes of ions and neutrals at the aperture in front of the plasma. The neutral cracking ratio was calculated using the measured intensities and the procedure described below.

The N₂ and N atom TOF peaks are not resolved in the experiments. Therefore, several contributions to the relative sensitivities of N and N₂ must be considered to get an accurate cracking ratio. Consider the measured N₂ and N intensities, $I_{N_2^+}$ and I_{N^+} , which are proportional to the true flux of N₂, f_{N_2} , and N, f_N , by the following relations:

$$I_{N_2^+} \propto \sigma_{N_2^+} s_{N_2} f_{N_2}, \quad (4)$$

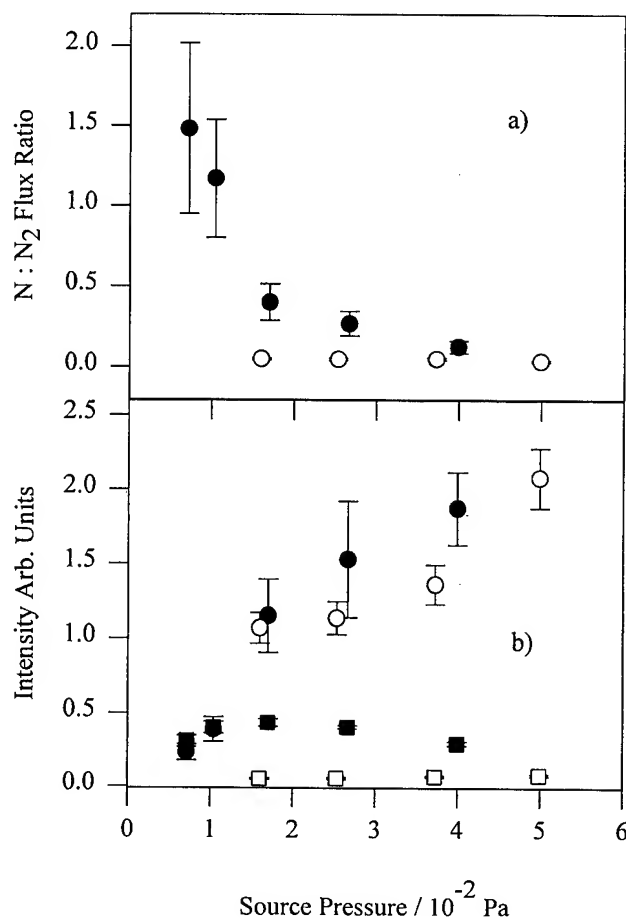


FIG. 6. Source gas cracking from N₂ ECR plasma vs source chamber pressure. Panel (a) shows corrected N:N₂ cracking ratio vs source pressure; (●) plasma on; (○) plasma off. (b) flux of neutrals emanating from the ECR source vs source pressure, (■) N, plasma ON; (□) N, plasma OFF; (●) N₂, plasma ON; (○) N₂, plasma OFF. Measurements are from the same plasma as data of Fig. 5.

$$I_{N^+} \propto \sigma_{N^+} s_N f_N + \sigma'_{N^+} s'_N f_{N_2}, \quad (5)$$

where σ_{N^+} is the cross section for electron impact ionization of N₂ to form the molecular ion, σ'_{N^+} is the cross section for electron impact dissociative ionization of N₂, σ_{N^+} is the cross section for the electron impact ionization of N atom, s_{N_2} and s_N are the QMS collection sensitivity factors for N₂⁺ and N⁺ formed *via* direct ionization of incident flux of N₂ and N, respectively, and s'_N is the collection sensitivity factor for N⁺ formed *via* dissociative ionization of N₂. Assuming an identical proportionality, which depends on the electron density and energy in the ionizer of the QMS, we solve these two equations for the ratio of incident flux of N and N₂ in terms of the measured intensities, cross sections and sensitivities. We arrive at Eq. (6) below:

$$\frac{f_N}{f_{N_2}} = \frac{I_{N^+} \sigma_{N_2^+} s_{N_2} - I_{N_2^+} \sigma'_{N^+} s_N}{I_{N_2^+} \sigma_{N^+} s_N}. \quad (6)$$

The cross sections used above are known for each of the three ionization processes for 70 eV electrons, which is the electron energy for the QMS ionizer used in this experiment. ($\sigma_{N_2^+} = 1.75 \times 10^{-16} \text{ cm}^2$; $\sigma'_{N_2^+} = 5 \times 10^{-17} \text{ cm}^2$; $\sigma_{N^+} = 1.55 \times 10^{-16} \text{ cm}^2$).³⁴ To calculate the true cracking ratio due to the plasma, f_N/f_{N_2} , from the measured intensities, the relative sensitivities, s_{N_2}/s_N and s'_N/s_N , must be determined.

The first of these, s_{N_2}/s_N , stems from the mass dependent probability of transmission through the QMS and can be estimated by a simple experiment where known fluxes of species, with different mass and known electron impact ionization cross sections, are measured with the QMS. To do this, three experiments are conducted where the source chamber was filled with each of Ne, Ar, and Kr gases, so that the densities of each gas measured by the QMS are equal. Analysis of the measured intensity of each gas, accounting for the known electron impact ionization cross sections,³⁴ gives the mass dependent transmission and detection efficiency of the QMS. From this data the relative sensitivity of the QMS to fluxes of N and N₂, $s_{N_2}/s_N = 0.82 \pm 0.07$, is determined. The relative sensitivity of the QMS to N⁺ formed via dissociative ionization of N₂ and N atom ionization, s'_N/s_N , is all that is still needed to determine the cracking ratio, f_N/f_{N_2} . The ratio, s'_N/s_N , is not equal to one since the collection efficiency of N⁺ produced via dissociative ionization of N₂ will be reduced due to the isotropic deposition of kinetic energy into the atomic fragments during the dissociative ionization process.³³

The determination of s'_N/s_N follows from the ability to determine s_N/s_{N_2} , described above, and the fact that the relative sensitivity, s'_N/s_{N_2} , can be determined from the measured cracking of thermal N₂. The intensity of N⁺ and N₂⁺ from cracking of thermal N₂ in the ionizer are related to the incident flux of N₂, f_{N_2} , by

$$I_{N_2^+} \propto \sigma_{N_2^+} s_{N_2^+} f_{N_2}, \quad (7)$$

$$I_{N^+} \propto \sigma'_{N^+} s'_{N^+} f_{N_2}, \quad (8)$$

where the definition of cross sections and sensitivities are defined above. It follows that

$$\frac{s'_N}{s_{N_2}} = \frac{I_{N^+} \sigma_{N_2^+}}{I_{N_2^+} \sigma'_{N^+}}. \quad (9)$$

With the cross sections and the measured intensity from thermal cracking we obtain $s'_N/s_{N_2} = 0.20 \pm 0.03$ and, hence, $s'_N/s_N = 0.17 \pm 0.04$. Now, with these relative sensitivities, cross sections and the measured N and N₂ intensities, the neutral cracking ratio is determined. Cracking ratios determined in this fashion are reported for a N₂ ECR plasma, as a function of source chamber pressure with a constant microwave power of 200 W and magnet current of 17 A and are shown in Fig. 6(a). The raw intensities of N and N₂ emanating from the ECR source, with the source on and off, are shown in Fig. 6(b). These intensities are not scaled by any of the corrections mentioned above.

The source gas cracking ratio, N:N₂, from Fig. 6(a), shows that the cracking ratio increases with decreasing pressure over the range measured and varies between 0.2 and 1.4. This is expected since the dissociation probability will be highest at the lower pressure due to higher electron energy. The cracking ratio reaches a maximum of 1.4 ± 0.5 at a pressure of $7.5 \times 10^{-3} \text{ Pa}$. Hence, at this pressure, 41% of the N₂ is cracked in the plasma. This is greater than the reported cracking of 12% determined via chemical titration by Meikle *et al.*,²⁸ however, those measurements are made after a considerably long path through a narrow tube, allowing a greater opportunity for N atom recombination on the tube surface.

Many applications of ECR's are dependent on the flux of reactive radicals which exit the source rather than the cracking ratio. As in the case of growing group III nitrides, one needs a high flux of N atoms to react with substrate materials. Figure 6(b) depicts the unscaled intensities of N and N₂ with the plasma on and off, from which the cracking ratios were calculated. The measured flux of N atoms increases with decreasing pressure down to $\sim 1.6 \times 10^{-2} \text{ Pa}$ and then the N atom flux decreases again as the pressure is lowered further. For a constant electron density and electron energy distribution function (EEDF) in the source, one would expect the N atom production to decrease with decreasing N₂ density. However, if the EEDF increases in energy, as is suggested from measurements by Hopwood *et al.*,²² the EEDF may overlap with a resonance that would increase the rate of N atom production via electron impact dissociation of N₂ at lower pressures. No attempt has been made to measure the electron density or energy concurrently with the TOF measurements. This behavior could have a significant effect on N atom ECR source applications.

Note that the thermal measurements in Fig. 6 are for gases introduced through the source, which results in noneffusive flow as eluded to above, where the thermal Ar mean translational energy is measured for different inlets. This directed noneffusive flow leads to a nonzero beam intensity, when linearly extrapolated to zero source chamber pressure. For purely effusive beams, where gas is introduced through a source not on the line-of-sight with the beam axis, the zero pressure beam intensity extrapolates to zero within experimental error. Direct comparison of intensities in Fig. 6(b) should be avoided, since data with the plasma on and off was collected many days apart, allowing the condition of the detector to drift.

C. Chlorine

1. Neutral mean kinetic energies

Mean kinetic energies for neutral atomic and molecular chlorine from a pure chlorine ECR plasma are presented in Figs. 7(a) and 7(b), along with mean kinetic energies with the plasma off. Mean kinetic energies are determined from a fit to the experimental TOF data as described in detail for Ar. As is the case with the nitrogen data, Cl TOFs have contamination from Cl⁺ produced via electron impact dissociation of Cl₂ in the ionizer of the QMS. Therefore, the Cl

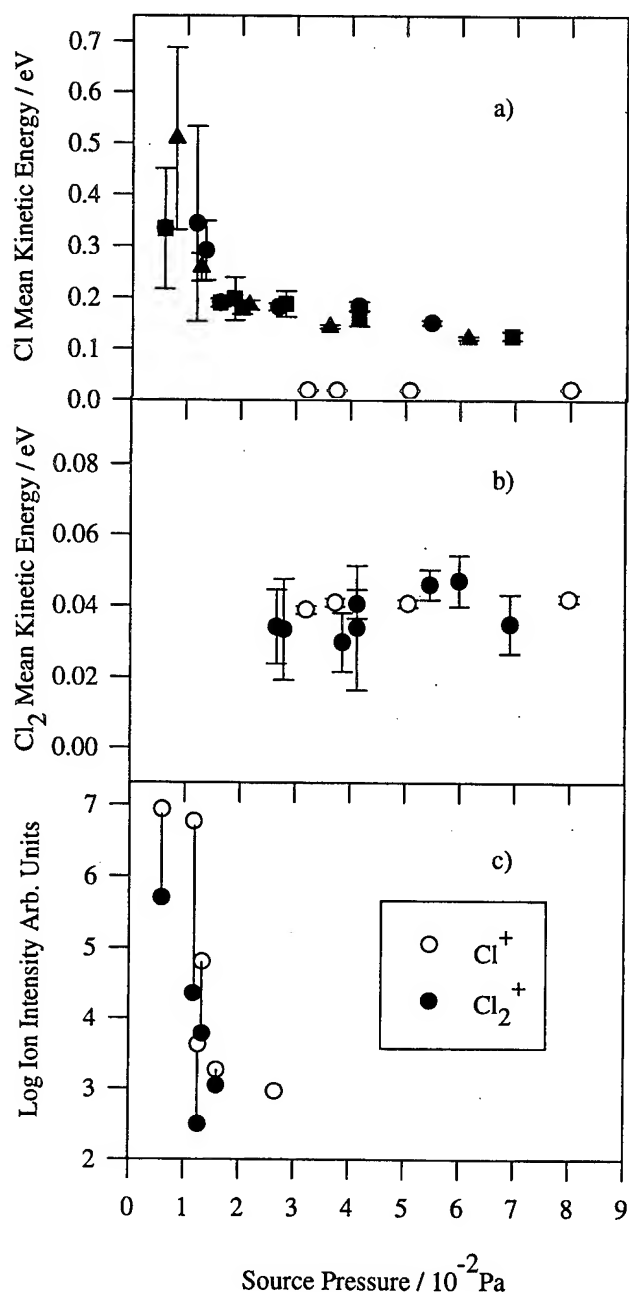


FIG. 7. Neutral mean kinetic energy vs source chamber pressure for (a) Cl; (b) Cl₂. ●, ■, and ▲ indicate data with plasma on taken on different days; ○ indicates data with the plasma off. (c) Ion flux intensity vs source chamber pressure for the same plasmas. Ion data from one series of runs is depicted. Microwave power of 200 W and magnet current of 17 A.

kinetic energies are slightly lower than if there are no contamination; however, the amount of Cl⁺ produced in the QMS compared to that produced in the ECR source is small, so this correction has been neglected. The Cl and Cl₂ data are taken immediately after one another, before changing plasma parameters to ensure identical plasma conditions for each species measurement. Figure 7(c) displays the Cl⁺ and Cl₂⁺ ion intensities emitted from the ECR source as a function of source pressure. These data are taken at the same time as one

of the runs of data pictured in Fig. 7(a) and 7(b). Data taken on other days are not reported since the detector efficiency changed with each day of chlorine experiments, due to chlorine attack on the electron multiplier in the QMS. However, the trends indicated in Fig. 7(c) are present in the chlorine ion data taken on different days. The ion data for Cl⁺ are reduced by a mass transmission efficiency factor, $s_{\text{Cl}}/s_{\text{Cl}_2} = 2.8 \pm 0.6$, which is determined via the calibration procedure mentioned above. There is no contamination of the Cl⁺ ion intensity from dissociative ionization of Cl₂, because the ionizer of the QMS is not on during ion data collection. Additionally, vertical lines are shown in the plot to group the Cl⁺ and Cl₂⁺ measurements taken from the same plasmas.

The Cl neutral mean kinetic energy data in Fig. 7(a) are similar to that for Ar and N atom, where the neutral mean kinetic energy rises with the increase in ion signal, seen in Fig. 7(c). However, the quality of the data at low pressure is poor due to low signal-to-noise in the TOF distributions. The mean kinetic energies of Cl and Cl⁺ within a chlorine ECR plasma were measured with Doppler spectroscopy and are also presented in Table I.^{14,19} The kinetic energies measured here (0.1–0.5 eV) are found to be much higher than those for neutral Cl measured by the Doppler method. In our measurements neutrals are produced through charge exchange with ions that pass through the sheath, which have greater kinetic energy, raising the mean kinetic energy of the neutrals exiting the source. The Cl₂ mean kinetic energies, shown in Fig. 7(b), are indistinguishable from the thermal Cl₂ data pictured with open symbols. The reason for this behavior is not clear, however, it may be that the only Cl₂ molecules that escape undissociated have not undergone collisions with the energetic species in the source and, therefore, are nearly thermal. Data for Cl₂ are only presented down to 3×10^{-2} Pa, below which the TOF distributions are obscured by the background. It is worth noting, that simulations by Ono *et al.* suggest that Cl₂⁺ is the major ion produced in Cl₂ ECR plasmas.²⁶ However, as seen in Fig. 7(c), where the vertical lines indicate measurements made of Cl⁺ and Cl₂⁺ from the same plasma, the Cl⁺ ion intensity is greater than the Cl₂⁺ ion intensity for all measurements.

2. Source gas cracking

Figure 8 shows the Cl:Cl₂ flux ratio as a function of source chamber pressure for constant microwave power and magnet current. The intensities of the neutrals are determined by integrating the area under the fits of the experimental TOF distributions. As with the nitrogen data, the transmission efficiency and electron impact ionization cross sections must be considered in calculating accurate cracking ratios from the measured ion intensities. The individual cross sections for the production of molecular Cl₂⁺ and Cl⁺ from the Cl₂ parent via electron impact ionization are not known; however, the total ionization cross section for Cl₂ is known for 70 eV electrons (7.1×10^{-16} cm² and 6.5×10^{-16} cm², respectively).^{35,36} The cross section for the electron impact ionization of Cl atom with 70 eV electron energy is also known (3.51×10^{-16} cm² and 3.47×10^{-16} cm²,

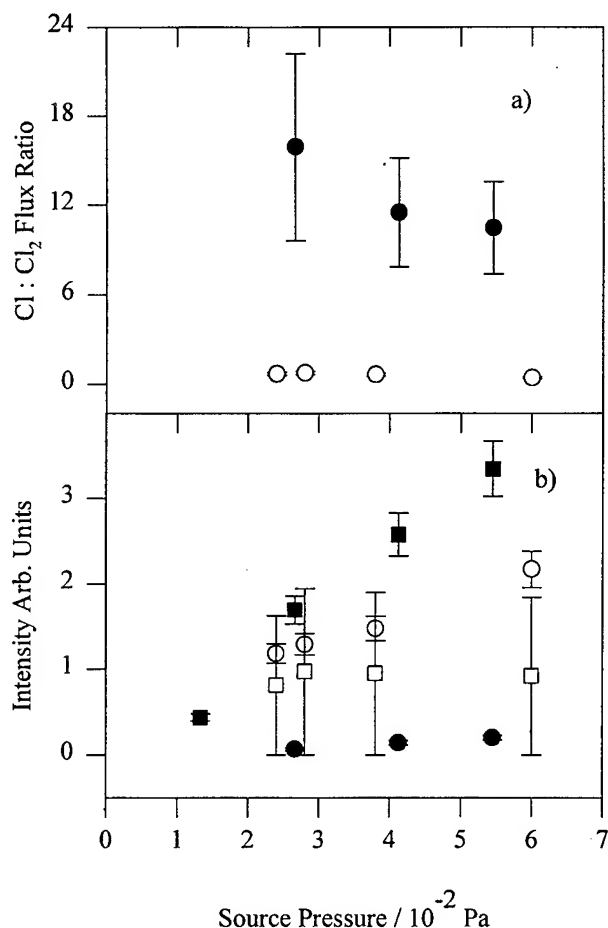


FIG. 8. Source gas cracking from Cl₂ ECR plasma vs source chamber pressure. Panel (a) shows corrected Cl:Cl₂ cracking ratio vs source pressure; (●) plasma on; (○) plasma off. (b) Flux of neutrals emanating from the ECR source vs source pressure, (■) Cl, plasma ON; (□) Cl, plasma OFF; (●) Cl₂, plasma ON; (○) Cl₂, plasma OFF. Measurements from the same plasma as data of Fig. 7.

respectively).^{37,38} The contribution to the measured Cl atom intensity from dissociative ionization of Cl₂ in the QMS is neglected, since this contribution is small compared to the flux of Cl atoms from the ECR source. The electron impact ionization cross section to form Cl₂⁺ from Cl₂, $\sigma_{Cl_2^+}$, is estimated from Cl:Cl₂, from the thermal Cl₂ cracking data, and the total electron impact ionization cross section for Cl₂ with 70 eV electron energy. The thermal cracking ratio, Cl:Cl₂, as measured with the QMS at 70 eV electron energy, is 0.6 ± 0.1 . The true cracking ratio will be higher, since the collection efficiency of Cl⁺ produced by dissociative ionization is less than that of Cl₂⁺ produced by electron impact of Cl₂, as discussed above for nitrogen. Using this cracking ratio, the estimated cross section for the electron impact ionization of Cl₂ to form Cl₂⁺, $\sigma_{Cl_2^+}$, is estimated to be $37\% \pm 6\%$ of the total ionization cross section for Cl₂. With these approximations, the ratio of the incident flux of Cl and Cl₂, f_{Cl}/f_{Cl_2} , may be expressed in terms of the measured intensities of Cl, I_{Cl^+} , and Cl₂, $I_{Cl_2^+}$, the cross sections, and the sensitivities by

$$\frac{f_{Cl}}{f_{Cl_2}} = \frac{I_{Cl^+} \sigma_{Cl_2^+} s_{Cl_2}}{I_{Cl_2^+} \sigma_{Cl^+} s_{Cl}}, \quad (10)$$

where s_{Cl}/s_{Cl_2} is the relative sensitivity reported above, and σ_{Cl^+} is the ionization cross section for ionization of Cl with 70 eV electrons. The cracking ratios in Fig. 8(a) are determined from the scaled ratio of the measured intensities as a function of source chamber pressure.

The Cl:Cl₂ flux ratio data for the Cl₂ ECR plasma shown in Fig. 8(a) display trends similar to those for the N₂ plasma shown in Fig. 6(a). The ratio increases from 10 ± 2 to 16 ± 6 with decreasing pressure, which is expected because the electron density increases with decreasing pressure and the EEDF increases to higher energy. Both of these will increase the dissociation probability at lower pressure. The degree of cracking of Cl₂ in this plasma far exceeds that of N₂. Here, at the lowest pressure where a Cl₂ TOF could be measured, the cracking of Cl₂ was 94%. This is contrary to the previously reported ratio of 1%–2%,²⁵ but is in agreement with a theoretical prediction for a non-ECR discharge.²⁷ The reason for this difference is unclear, however, the reactor geometries used in this study and that of Ono *et al.*²⁵ are quite different. The reason for different cracking behavior in N₂ and Cl₂ plasmas may be due to the much weaker Cl₂ bond. These results do show that the plasma behavior is very dependent on the nature of the source gas used and that it is difficult to draw general conclusions concerning ECR plasmas in different source gases. The differences are in part due to the differences in cross sections for electron impact dissociation. Additionally, plasma properties such as the electron density, and the EEDF for each plasma may be different and depend on the electronic states of the plasmas constituents.

Figure 8(b) shows the raw intensity of neutral Cl and Cl₂ as a function of source chamber pressure. Intensities were determined by integrating the area under the fits to the experimental TOF data. Note that as in the case of the thermal nitrogen, the thermal chlorine flow was introduced through the source and results in noneffusive behavior, which leads to nonzero beam intensity when linearly extrapolated to zero source chamber pressure. The intensities were not scaled by any of the factors mentioned in the cracking treatment. Here, unlike the case for nitrogen, the Cl flux from the plasma continuously increases with increasing source pressure. This is largely due to the very high degree of cracking in the plasma. Slight changes in the plasma, affecting the rate of cracking, would not be sufficient to increase the flux of Cl atom with the decreasing stock of Cl₂ source gas.

It is interesting to consider how the behavior of this Cl₂ ECR plasma might affect chlorine etching applications, especially since there are many competing factors. Ion intensity, with its undesirable charging effects, is nonetheless a major element in determining etch rates. Generally, the greater the ion flux to the surface, the greater the etch rate; however, loading of the neutral reactive species will also be important. If there are not enough reactive species, the rate of the etching reaction will not occur at a high rate even in the

presence of a high ion flux. Therefore, there will be a competition in the selection of plasma parameters between the reactive particle loading and ion flux requirements. Lower pressure favors ion flux but limits flux of reactive radicals, and higher pressure limits the ion flux but increases the flux of reactive species.

IV. CONCLUSIONS

Modulated beam TOF measurements are employed to study Ar, N₂, and Cl₂ ECR plasmas. Neutral mean kinetic energy, ion flux, and source gas cracking information are extracted from measured TOF distributions. Mean neutral kinetic energies vary between 0.04 and 0.45 eV. Enhancement in neutral mean kinetic energy shows a strong correlation with ion flux out of the source for all source gases studied. At pressures below 2.0×10^{-2} Pa, ion flux sharply increases causing the neutral mean kinetic energy to increase rapidly. This effect is much stronger in atomic neutrals as compared to molecular neutrals. This effect is not observed for Cl₂, since the molecular signal in the low pressure regime was too low. Source gas cracking is also examined as a function of source pressure. The cracking ratios of both N₂ and Cl₂ are found to increase with decreasing pressure; however, the Cl₂ cracking ratio is an order of magnitude greater than that of the N₂ cracking ratio. The fluxes of atomic and molecular species were also measured. The flux of N atoms emanating from the ECR plasma is found to first increase, and then decrease with increasing source pressure. This behavior is not observed for chlorine where the Cl flux increased, continuously with increasing source pressure.

The technique of modulated beam TOF-QMS is demonstrated as a reliable tool for monitoring the plasma environment, to reveal the behavior of single source gas plasmas in terms of neutral kinetic energies, and relative neutral and ionic fluxes downstream from the plasma. This is a robust technique which can be extended to measure plasmas of complex source gas mixtures, that would be difficult to measure with optical methods. The kinetic energy results are consistent with the many previous studies in which optical Doppler width measurements were employed to determine neutral kinetic energies. Modulated beam mass spectrometry is not dependent on the electronic nature of the particle in question, and can be extended to the measurement of velocity distributions of atomic or molecular species, where the lack of a metastable state would preclude the use of Doppler spectroscopy to measure mean kinetic energies. Additionally, these measurements are insensitive to any correlation that may exist between kinetic energies and electronic states of particles in the plasma. Additional studies, coupling Langmuir probe measurements with neutral and ion flux measurements could provide a very complete picture of the ECR plasma environment, to further the understanding of these complex interdependent systems.

ACKNOWLEDGMENTS

The authors acknowledge the support of this work by the U.S. Army Research Office, Physics Division. Additional equipment and support were provided by the National Science Foundation and the National Institute of Standards and Technology.

- ¹J. Asmussen, *J. Vac. Sci. Technol. A* **7**, 883 (1989).
- ²J. Grenier, *Solid State Technol.* **31**, 67 (1988).
- ³V. M. Donnelly and D. L. Flamm, *Solid State Technol.* **24**, 161 (1981).
- ⁴H. F. Winters, J. W. Coburn, and T. J. Chuang, *J. Vac. Sci. Technol. B* **1**, 469 (1983).
- ⁵A. M. Voshchenkov, *J. Vac. Sci. Technol. A* **11**, 1211 (1993).
- ⁶S. R. Leone, *Jpn. J. Appl. Phys., Part 1* **34**, 2073 (1995).
- ⁷M. A. Lieberman and A. J. Lichtenberg, *Principles of Plasma Discharges and Materials Processing* (Wiley, New York, 1994).
- ⁸H. F. Winters and J. W. Coburn, *Surf. Sci. Rep.* **14**, 161 (1992).
- ⁹G. King, F. C. Sze, P. Mak, T. A. Grotjohn, and J. Asmussen, *J. Vac. Sci. Technol. A* **10**, 1265 (1992).
- ¹⁰R. A. Stern and J. A. Johnson, *Phys. Rev. Lett.* **34**, 1548 (1975).
- ¹¹E. A. Den Hartog, H. Persing, and R. C. Woods, *Appl. Phys. Lett.* **57**, 661 (1990).
- ¹²T. Nakano, N. Sadeghi, and R. A. Gottscho, *Appl. Phys. Lett.* **58**, 458 (1991).
- ¹³N. Sadeghi, T. Nakano, D. J. Trevor, and R. A. Gottscho, *J. Appl. Phys.* **70**, 2552 (1991).
- ¹⁴T. Nakano, N. Sadeghi, D. J. Trevor, R. A. Gottscho, and R. W. Boswell, *J. Appl. Phys.* **72**, 3384 (1992).
- ¹⁵J. Hopwood and J. Asmussen, *Appl. Phys. Lett.* **58**, 2473 (1991).
- ¹⁶D. V. Tsu, R. T. Yong, S. R. Ovshinsky, C. C. Klepper, and L. A. Berry, *J. Vac. Sci. Technol. A* **13**, 935 (1995).
- ¹⁷J. S. McKillop, J. C. Forster, and W. M. Holber, *J. Vac. Sci. Technol. A* **7**, 908 (1989).
- ¹⁸J. S. McKillop, J. C. Forster, and W. H. Holber, *Appl. Phys. Lett.* **55**, 30 (1989).
- ¹⁹J. Wormhoudt, A. C. Stanton, A. D. Richards, and H. H. Sawin, *J. Appl. Phys.* **61**, 142 (1987).
- ²⁰S. M. Rossnagel, S. R. Whitehair, C. R. Guarnieri, and J. J. Cuomo, *J. Vac. Sci. Technol. A* **8**, 3113 (1990).
- ²¹T. Tsuchizawa, Y. Jin, and S. Matsuo, *Appl. Phys. Lett.* **69**, 149 (1996).
- ²²J. Hopwood, D. K. Reinhard, and J. Asmussen, *J. Vac. Sci. Technol. A* **8**, 3103 (1990).
- ²³M. D. Kilgore, H. M. Wu, and D. B. Graves, *J. Vac. Sci. Technol. B* **12**, 494 (1994).
- ²⁴T. J. Sommerer and M. J. Kushner, *J. Appl. Phys.* **70**, 1240 (1991).
- ²⁵K. Ono, M. Tuda, K. Nishikawa, T. Oomori, and K. Namba, *Jpn. J. Appl. Phys. Part 1* **33**, 4424 (1994).
- ²⁶K. Ono, M. Tuda, H. Ootera, and T. Oomori, *Pure Appl. Chem.* **66**, 1327 (1994).
- ²⁷S. C. Deshmukh and D. J. Economou, *J. Appl. Phys.* **72**, 4597 (1992).
- ²⁸S. Meikle and Y. Hatanaka, *Appl. Phys. Lett.* **54**, 1648 (1989).
- ²⁹R. P. Vaudo, J. P. Cook, Jr., and J. F. Schetzina, *J. Vac. Sci. Technol. B* **12**, 1232 (1994).
- ³⁰*Atomic and Molecular Beam Methods*, edited by G. Scoles (Oxford University Press, Oxford, 1988), Vol. 1.
- ³¹W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, New York, 1994).
- ³²W. H. Cramer, *J. Chem. Phys.* **30**, 641 (1959).
- ³³J. M. Ajello and M. Ciocca, *J. Geophys. Res.* **101**, 18 953 (1996).
- ³⁴*Ionization Cross Sections of Atoms and Ions by Electron Impact*, edited by H. Tawara, T. Kato, and M. Ohnishi (Institute of Plasma Physics, Nagoya, Japan, 1985).
- ³⁵F. A. Stevie and M. J. Vasile, *J. Chem. Phys.* **74**, 5160 (1981).
- ³⁶R. E. Center and A. Mandl, *J. Chem. Phys.* **57**, 4104 (1972).
- ³⁷R. S. Freund, R. C. Wetzel, R. J. Shul, and T. R. Hayes, *Phys. Rev. A* **41**, 3575 (1990).
- ³⁸T. R. Hayes, R. C. Wetzel, and R. S. Freund, *Phys. Rev. A* **35**, 578 (1987).

Dry etch damage in GaAs metal-semiconductor field-effect transistors exposed to inductively coupled plasma and electron cyclotron resonance Ar plasmas

F. Ren

Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974

J. W. Lee, C. R. Abernathy, and S. J. Pearton^{a)}

Department of Materials Science and Engineering, University of Florida, Gainesville, Florida 32611

C. Constantine and C. Barratt

Plasma Therm IP, St. Petersburg, Florida 33716

R. J. Shul

Sandia National Laboratories, Albuquerque, New Mexico 87185

(Received 4 December 1996; accepted 11 April 1996)

The effects of Ar plasma exposure on transconductance, channel sheet resistance, output resistance, and gate contact ideality factor of GaAs metal-semiconductor field-effect transistors (MESFETs) were investigated using two different high-density plasma sources, namely inductively coupled plasma and electron resonance plasma. Ion-induced damage is found to be reduced at moderate source powers (~ 200 W) because of the reduction in cathode dc self-bias and hence ion energy, but at higher source powers the increase in ion flux produces significant deterioration of the device performance. Careful attention must be paid to both ion flux and ion energy in order to minimize ion-induced damage. Due to their relatively low channel doping levels, MESFETs are found to be more sensitive to plasma damage than devices with very heavily doped component layers such as heterojunction bipolar transistors. © 1997 American Vacuum Society. [S0734-211X(97)03304-0]

I. INTRODUCTION

The GaAs metal-semiconductor field-transistor (MESFET) is still the workhorse of III-V electronic device technology, with wide-ranging applications in wireless communication, collision avoidance radar systems, and intelligent cruise control for automobiles, digital signal processing, and power microwave systems.¹⁻⁵ The majority of these devices are fabricated by direct ion implantation into semi-insulating GaAs substrates, with a selective area channel (for n -layer formation) and selective area source/drain (for n^+ contact formation) implantation process. To simplify the lithography requirements it is also common to form sequential n^+/n layers by implantation or epitaxy, and then etch down to the channel layer prior to gate metal deposition.² Finally, use of SiN_x surface passivation layers deposited by plasma-enhanced chemical vapor deposition is common for achieving stable long-term surface characteristics. These films are generally patterned in areas requiring interconnects (source, drain, and gate pads) by dry etching with a high selectivity for nitride over GaAs.⁶⁻¹⁰

Based on the ubiquity of dry etching in GaAs MESFET processing,^{11,12} an understanding of possible ion-induced damage occurring during these steps is highly desirable. This is particularly necessary with the recent emphasis on high-density plasma sources that operate at ion densities approximately two orders of magnitude higher than in conventional capacitively coupled reactive ion etching (RIE) systems. The basic configuration of these new single-wafer reactors is

plasma generation in an (almost) remote high-density plasma source, with ion energies of 20–30 eV, ion densities of $>10^{11} \text{ cm}^{-3}$, and separate rf biasing of the sample position to control the incident ion energy.¹³⁻¹⁹ This additional acceleration of the ions may be necessary to achieve the requisite etch rates and anisotropy. The high ion fluxes in these new etch tools have made possible improved pattern transfer processes for materials with involatile etch products such as InP ,^{20,21} and high bond strengths, such as GaN and AlN .²² However, there has been little study into the damage induced in device structures for different plasma conditions. One way to examine this is to employ nonetching plasma chemistries, such as Ar, which simulates the ion bombardment component received by a device during an etch process.^{6,8,10,11} Argon is a particularly relevant choice since it is generally added to plasma recipes in high-density etch tools to facilitate discharge ignition at the low pressures (~ 1 mTorr) that are typically used. Since the sputter rates with pure Ar plasmas are much lower than the etch rates with a chemistry such as Cl_2/Ar , the ion-induced damage is able to accumulate without the accompanying material removal of a real etch process. Thus, the study of Ar-induced plasma change represents a worst case scenario and provides a valuable upper bound on the amount of device degradation expected.

We have commenced a systematic study of damage induced in MESFETs, high electron mobility transistors, and heterojunction bipolar transistors during various dry processes. In this article we compare the effects of Ar plasma exposure on the performance of GaAs MESFETs using two different types of high-density sources, namely inductively

^{a)}Electronic mail: spear@mse.ufl.edu

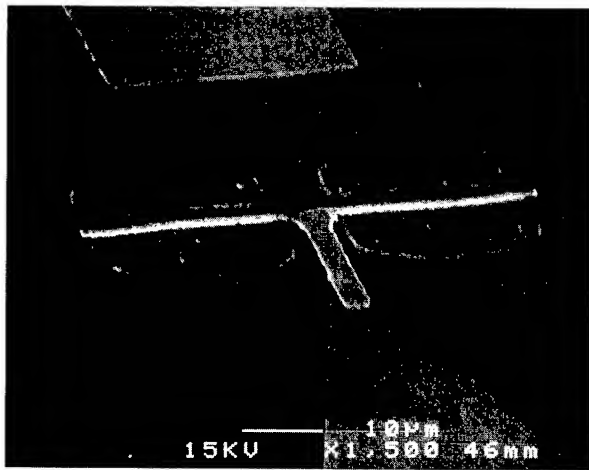


FIG. 1. Scanning electron microscope micrograph of MESFET.

coupled plasma (ICP) and electron cyclotron resonance (ECR). Most work to date on developing etch processes for III-V materials using high-density plasmas has been done with ECR sources that operate at microwave frequencies. However, the simpler design, potential for easier scale-up to large wafer diameters and more mature matching network technology available for rf ICP sources makes them an attractive choice for manufacturing applications.

II. EXPERIMENT

The MESFETs consisted of 4000 Å of Si-doped n -type ($n = 2 \times 10^{17} \text{ cm}^{-3}$) epitaxial GaAs grown on semi-insulating GaAs (100) substrates. AuGeNi source/drain contacts were patterned by lift-off and alloyed at 420 °C for 30 s, while TiPtAu gate contacts (1 μm gate length) were also patterned by lift-off. The gates are recessed into the channel by ~800 Å. Device isolation was achieved by Ar^+ ion milling (500 eV, 10 °C, vertical incidence), followed by a wet etch clean-up ($\text{NH}_4\text{OH}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$). A scanning electron micrograph of a completed device is shown in Fig. 1.

The devices were exposed to ICP or ECR Ar plasmas with the gates in place, to measure the effect of damage induced between the source-gate and gate-drain regions. The unpassivated MESFET is used as a vehicle for characterizing plasma damage—our conditions would not be part of a normal process. The ICP system was a Plasma-Therm 790 series reactor in which the disk geometry source was operated at 2 MHz and powers of 0–1000 W. The He backside-cooled sample chuck was biased with 13.56 MHz power (10–300 W) and produced dc self-biases of 0 to –340 V. Process pressure was held constant at 2 mTorr, with 20 sccm Ar injected into the ICP source through an electronic mass flow controller (MFC). The ECR system was a Plasma-Therm SLR 770 reactor in which the low profile (single magnet) Astex 4400 6 in. ϕ source was operated at 2.45 GHz and powers of 0–1000 W (the upper magnet current was 170 A and the lower magnet current was 40 A). The He backside-cooled sample chuck was biased with 13.56 MHz power (10–200 W) and produced dc self-biases of –25 to

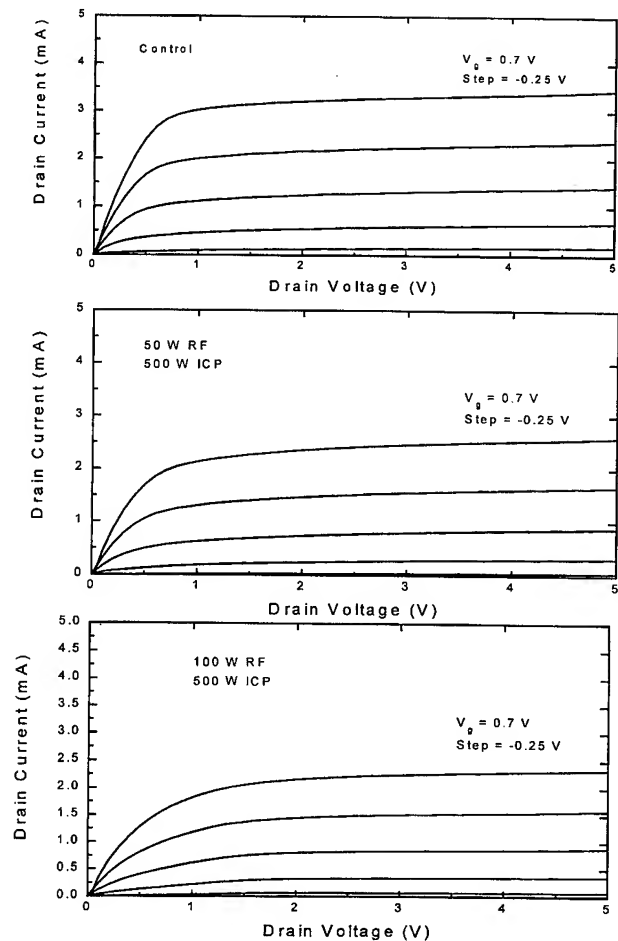


FIG. 2. $I_{DS}-V_{DS}$ characteristics for MESFETs before (top) or after ICP Ar plasma exposure (500 W source power, 15 s) at either 50 W (center) or 100 W (bottom) rf chuck power.

–340 V. Process pressure was again held constant at 2 mTorr with 20 sccm of Ar injected into the ECR source through a MFC. In both systems the source-sample distance was ~10 cm and plasma exposures were 15 s in duration. Loss of material by sputtering was <400 Å in all cases. Device characterization consisted of measuring the source-drain current voltage ($I_{DS}-V_{DS}$) characteristics as a function of gate voltage, the channel layer sheet resistance from transmission line patterns ($100 \times 100 \mu\text{m}^2$ AuGeNi contact pads) that are an integral part of the device maskset, the ideality factor n of the gate contact under forward bias, the device output resistance, and finally its transconductance, g_m . A key point to remember is that the ECR system develops higher dc self-biases than the ICP for a given rf chuck power because of the smaller electrode diameter (4 in. vs 8 in.).

III. RESULTS AND DISCUSSION

Figure 2 shows some typical $I_{DS}-V_{DS}$ characteristics for MESFETs before (top) and after (center and bottom) exposure to 500 W ICP plasmas with two different rf chuck powers. The decrease in saturation current is obvious in the plasma-exposed devices and is more pronounced at higher rf

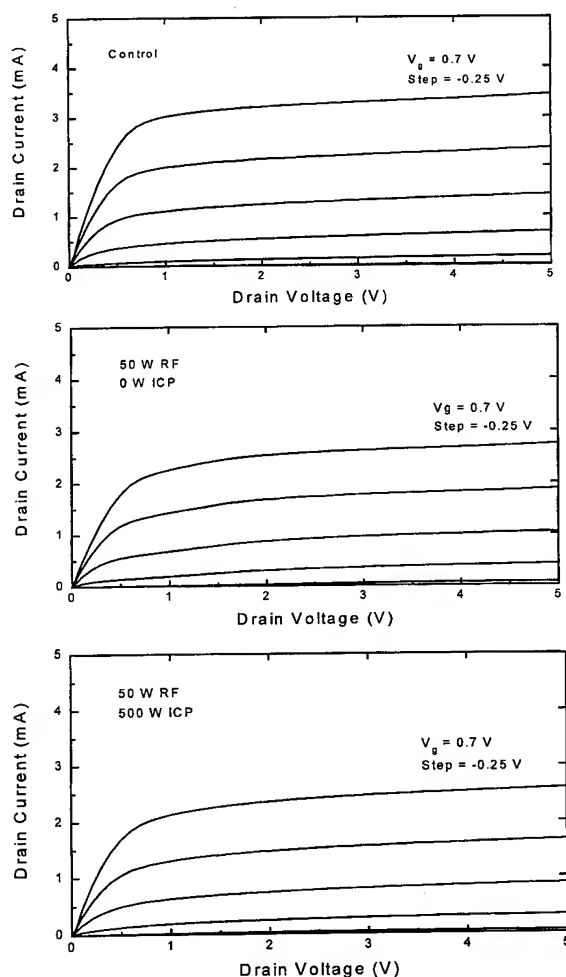


FIG. 3. $I_{DS}-V_{DS}$ characteristics for MESFETs before (top) and after ICP Ar plasma exposure (50 W rf chuck power, 15 s) at either 0 W (center) or 500 W (bottom) source power.

chuck powers. A comparison with the results in Figure 3 shows that this latter power is a stronger influence than ICP source power, at least to 100 W chuck power and 500 W source power. As seen in Figs. 2 and 3, saturated current decreases with increasing rf power and source power, but there is a stronger dependence on increasing rf power.

Similar data for MESFETs exposed to ECR Ar plasmas is shown for different rf chuck powers in Fig. 4 and for different ECR source powers in Fig. 5. In this case, both powers have a strong influence on device performance and this is related to the differences in ion energy between the two systems.

The high-density source power basically controls the ion flux incident on the sample position, while the chuck power controls ion energy. The effects of these two parameters were systematically investigated in both tools.

Figure 6 shows MESFET channel sheet resistance and gate contact ideality factor for devices exposed to ICP Ar discharges at fixed rf power (50 W), as a function of ICP source power. Note that the cathode dc self-bias is suppressed as source power is increased, even though ion flux

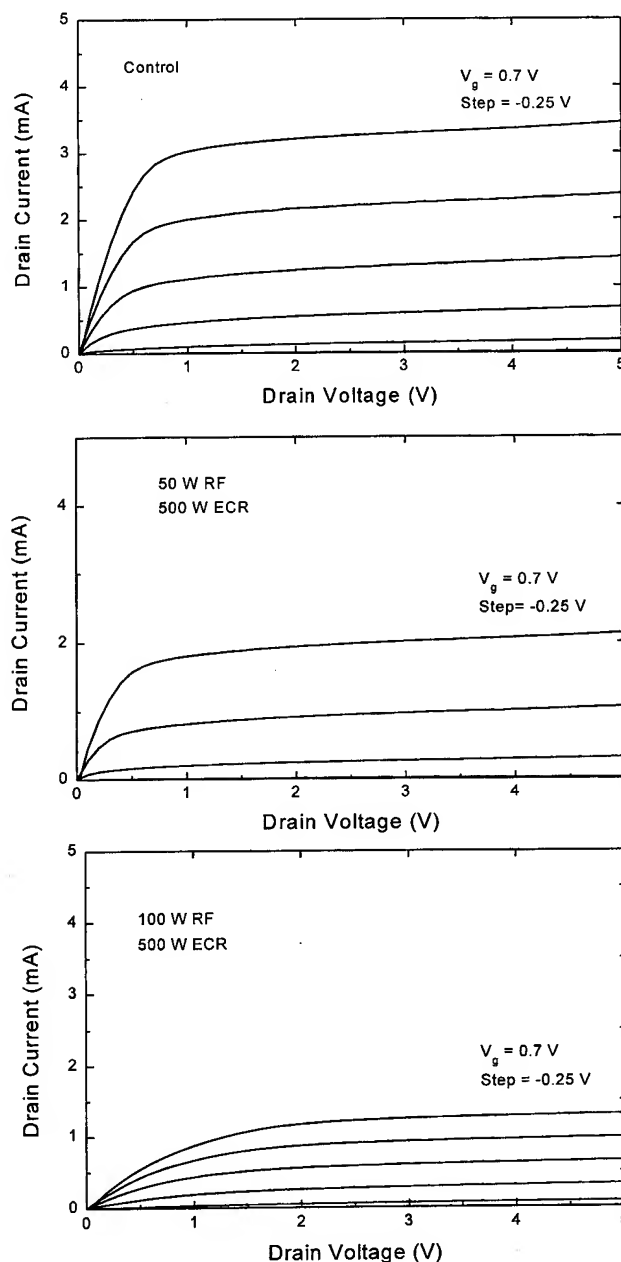


FIG. 4. $I_{DS}-V_{DS}$ characteristics for MESFETs before (top) and after ECR Ar plasma exposure (500 W source power, 15 s) at either 50 W (center) or 100 W (bottom) rf chuck power.

will increase. This is due to the higher plasma density under these conditions. The change in channel sheet resistance initially decreases as the ion energy decreases because there are fewer deep traps created per incident ion, but then increases as ion flux increases at higher ICP powers. Even when the dc self-bias is completely suppressed there will still be an ion energy of 20–30 eV because of the positive plasma potential relative to the sample position.¹⁶ The sheet resistance of the channel layer will be increased if the incident energetic ions create deep electron traps found by diffusion of point defects from the near-surface into the channel region that remove carriers from the conduction process. This is the same tech-

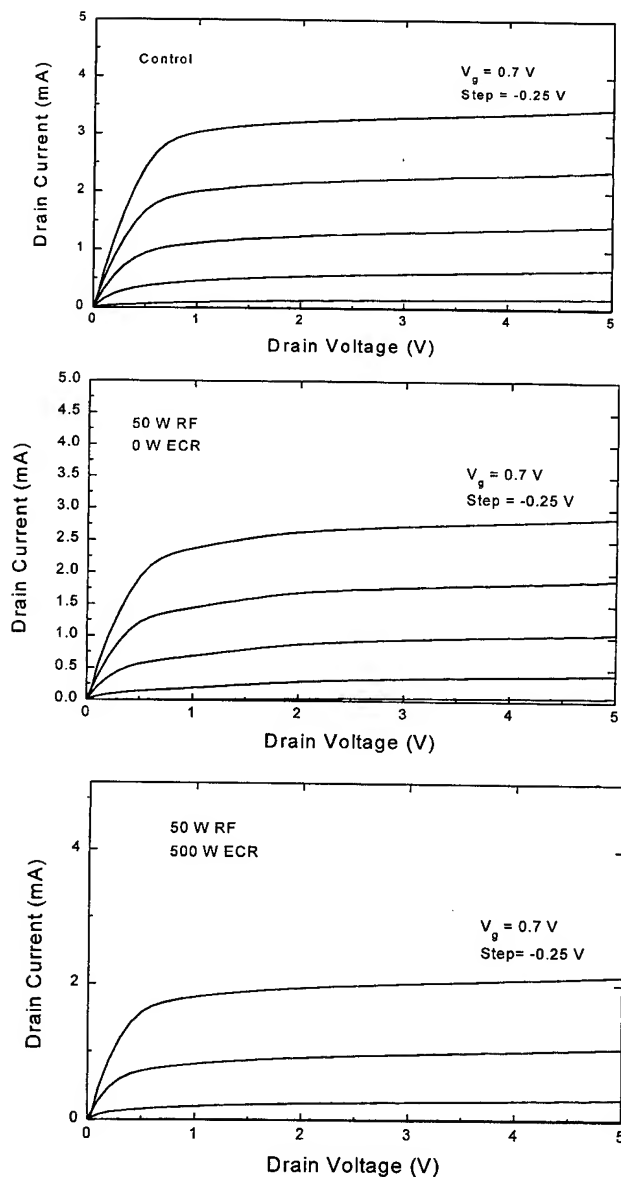


FIG. 5. $I_{DS}-V_{DS}$ characteristics for MESFETs before (top) and after ECR plasma exposure (50 W rf chuck power, 15 s) at either 0 W (center) or 500 W (bottom) source power.

nique used to create high resistance regions for electrical device isolation.²³ The ideality factor is unchanged under all conditions, indicating that surface damage around the periphery of the contact that might induce generation-recombination currents is not significant.²⁴

The effects of ICP source power on g_m and output resistance are shown in Fig. 7 for fixed rf power (50 W) plasma exposures. The transconductance is reduced by ~25% at low source powers (≤ 200 W) and decreases further as ion flux is increased. These changes are in good correlation with the behavior of the channel layer sheet resistance. The device output resistance decreases initially as ion flux increases (to ~500 W ICP source power), but improves somewhat thereafter as ion energy is reduced. The results in Figs. 6 and 7 suggest that moderate ICP source powers (200–500 W) pro-

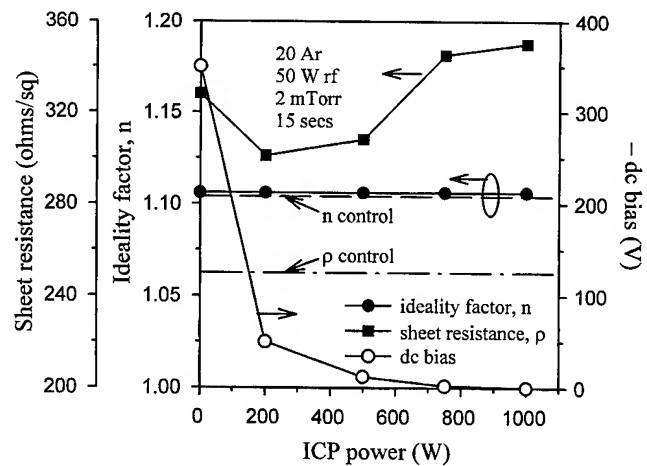


FIG. 6. Channel layer sheet resistance, gate control ideality factor, and cathode dc self-bias as functions of ICP source power for fixed rf chuck power (50 W).

duce the least device degradation, and emphasize the trade-off at fixed rf chuck power between lower ion energy (which occurs at high source power) and lower ion flux (which occurs at low source power).

MESFETs exposed to ECR Ar plasmas as a function of source power showed similar behavior to those treated in ICP discharges. Figure 8 shows the effect of increasing ECR microwave power on channel layer sheet resistance and gate contact ideality factor. Once again the dc self-bias is suppressed as the source power is increased, but remains at ~75 V even at 1000 W microwave power. This is most likely a result of the different electrode diameters (4 in. for the ECR tool, 8 in. for the ICP tool) and the resulting higher power per unit area in the ECR case. The sheet resistance is again least affected at moderate source power, and increases at higher powers in correlation with the higher ion flux incident on the MESFET. The gate diode ideality factor shows no measurable change with microwave power.

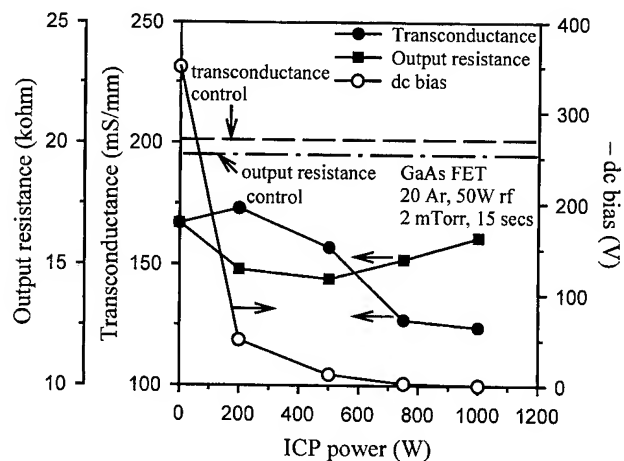


FIG. 7. Transconductance, output resistance, and cathode dc self-bias as functions of ICP source power for fixed rf chuck power (50 W).

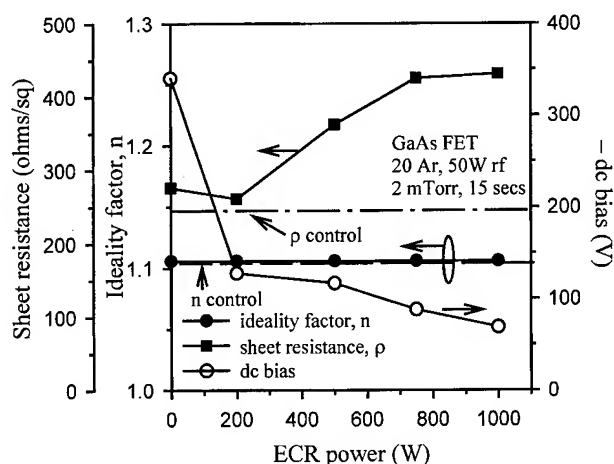


FIG. 8. Channel layer sheet resistance, gate contact ideality factor, and cathode dc self-bias as functions of ECR source power for fixed rf chuck power (50 W).

The effects of ECR source power on g_m and output resistance are shown in Fig. 9. Transconductance is least affected at source powers around 200 W, and decreases at higher powers indicating that ion flux is the dominant factor in degrading the device performance under these conditions. By contrast output resistance improves slightly at higher ECR source powers, and tracks the decrease in ion energy, as was the case with the ICP exposures.

As was seen in the $I_{DS} - V_{DS}$ characteristics, the rf chuck power has a strong influence on device properties of plasma exposed MESFETs. Figure 10 shows the effect of rf chuck power on the channel layer sheet resistance and ideality factor of MESFETs exposed to ICP discharges at fixed source power (500 W). The sheet resistance increases in an almost linear manner above ~50 W rf chuck power and is approximately a factor of 4 larger than the control value for devices exposed at 300 W.

In order to understand the relationship between the change in device characteristics, ion damage, and ion flux,

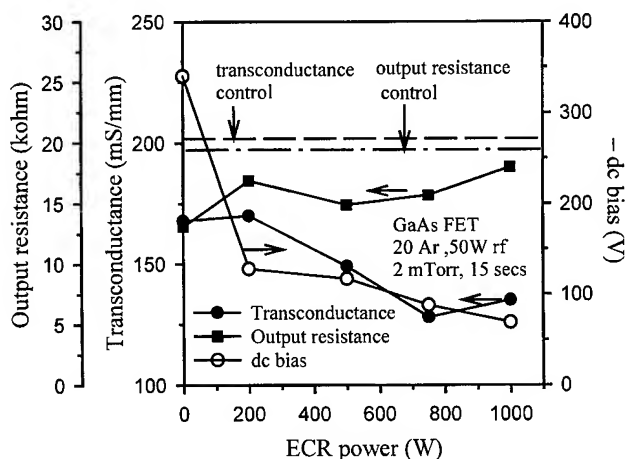


FIG. 9. Transconductance, output resistance, and cathode dc self-bias as functions of ECR source power for fixed rf chuck power (50 W).

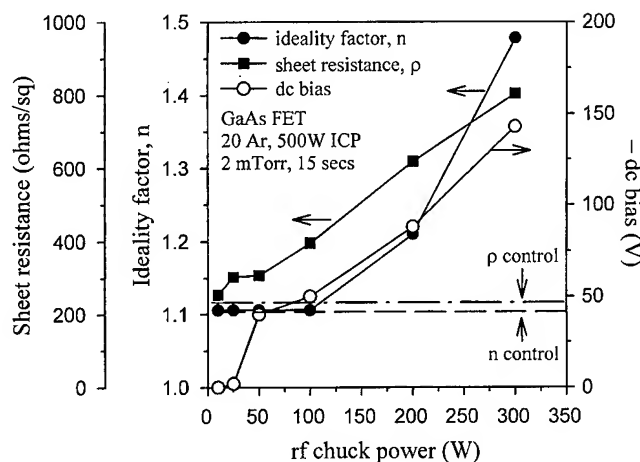


FIG. 10. Channel layer sheet resistance, gate contact ideality factor, and cathode dc self-bias as functions of rf chuck power for fixed ICP source power (500 W).

some calculations were performed. The number of defects created by impinging ions N_d is given by the Kinchin-Pease formula²⁵

$$N_d = E_n / 2E_d, \quad (1)$$

where E_n is the energy lost by nuclear stopping and E_d is the energy threshold for displacement of lattice atoms. If the self-bias at 300 W is -140 V, and the plasma potential adds another ~30 V, then the "Transport of Ions is Matter,"²⁶ ion stopping code predicts that approximately 15% of this 170 eV will go into nuclear stopping. Assuming E_d is ~25 eV for a semiconductor, then each ion will create approximately one displacement defect. The ion flux is of order 2×10^{17} ions $\text{cm}^{-2} \text{s}^{-1}$ for a 500 W source power, 300 W rf chuck power condition, and thus $\sim 3 \times 10^{18}$ defects cm^{-2} would be created in 15 s. However, in ion implantation experiments it is well established that defect recombination and annihilation greatly reduce the final amount of damage, and in our case the surface clearly does not become amorphous as judged by its reflectivity, and thus the number of stable defects remaining is $\ll 0.1\%$ of those created. Indeed the number of defects remaining must be comparable to the doping density, based on the sheet resistance increase, and thus averaging over the channel thickness (which is a gross simplification because the damage will be concentrated near the surface), one obtains an areal defect density of $\leq 5 \times 10^{13} \text{ cm}^{-2}$, i.e., only 1 in 6×10^4 defects is stable. There are several other options; the Kinchin-Pease formula may grossly overestimate defect production at these low energies, or many defects may be diffusing into the material to depths where they are electrically unimportant. This area has recently been reviewed by Hu *et al.*²⁷ and Murad *et al.*²⁸ Note also in Fig. 10 that the gate contact ideality factor rises above 100 W rf chuck power, indicating the onset of severe damage creation around the contact periphery. From these results the rf chuck power should be maintained below 50 W.

The influence of chuck power on g_m and output resistance is shown in Fig. 11 for devices exposed to ICP Ar plasmas at

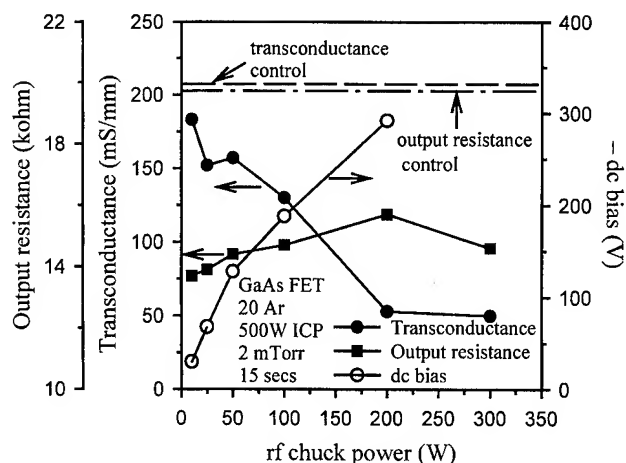


FIG. 11. Transconductance, output resistance, and cathode dc self-bias as functions of rf chuck power for fixed ICP source power (500 W).

fixed source power (500 W). Transconductance falls rapidly above 50 W chuck power in correlation with the sheet resistance and the increasing ion energy, whereas output resistance is less affected by ion energy, but is well below the control value under all conditions. The combined results of Figs. 7 and 10 clearly show that both rf chuck power and ICP source power should be kept to moderate values to minimize ion-induced damage to GaAs MESFETs. For dielectric etching in SF_6/Ar or CF_4/O_2 , ICP reactors offer the advantage over conventional RIE of lower ion energies due to the suppression of cathode dc self-bias as source power is increased.

The influence of rf chuck power on channel layer sheet resistance and gate contact ideality factor for ECR exposed devices at fixed source power is shown in Fig. 12. As was the case with the ICP exposed MESFETs, the sheet resistance of the channel increases rapidly above ~50 W rf chuck power. The ultimate change in this sheet resistance is worse than for the ICP reactor because of the higher dc bias for a given rf chuck power. The gate contact ideality factor also

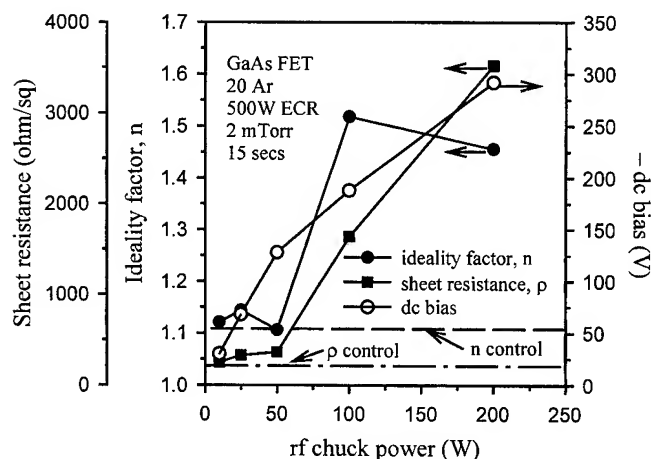


FIG. 12. Channel layer sheet resistance, gate contact ideality factor, and cathode dc self-bias as functions of rf chuck power for fixed ECR source power (500 W).

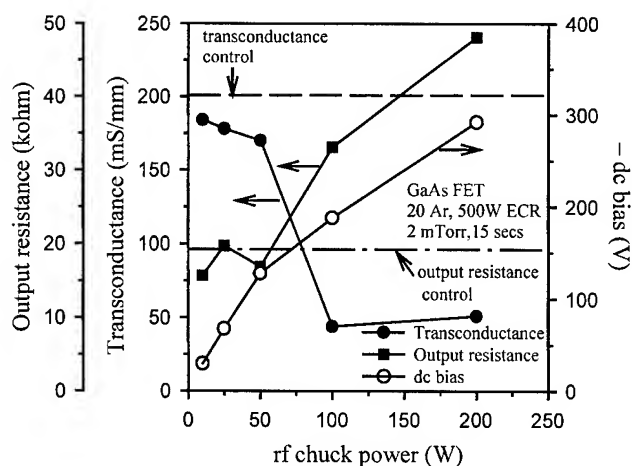


FIG. 13. Transconductance, output resistance, and cathode dc self-bias as functions of rf chuck power for fixed ECR source power (500 W).

increases rapidly above a chuck power of ~50 W, corresponding to the introduction of recombination centers around the periphery of this contact. The gate metal itself protects the region underneath it from ion-induced damage and thus the increase in ideality factor originates from degradation of the GaAs around the boundary of the contact.

Figure 13 shows the effect of rf chuck power on g_m and output resistance of MESFETs exposed to 500 W microwave power ECR discharges. Once again there is a rapid decrease in g_m above ~50 W rf chuck power as was the case with the ICP samples. The main difference is that for the ECR exposed devices the output resistance now increases well above the control value, which is related to the higher ion energies encountered in the ECR tool. These conditions clearly produce unacceptable changes in MESFET characteristics.

IV. SUMMARY AND CONCLUSIONS

There are several key results that emerge from this study:

(i) Both ion energy and ion flux contribute to device degradation in MESFETs exposed to high density plasmas. It is widely understood that dc self-bias should be minimized (consistent with achieving the necessary etch anisotropy) during dry etching, but high source powers can induce significant changes in device operating characteristics even at low ion energies.

(ii) Moderate ICP or ECR source powers produce the least change in MESFET performance because of the decrease in cathode self-bias as ion density increases. The ion-induced damage is less under these conditions than with RIE, and there should be an associated advantage of higher etch rates with the high density reactors.

(iii) In our reactors the ICP tends to produce slightly less damage to MESFETs than the ECR tool at similar source powers, which we believe is accounted for by the slightly lower dc bias for a given rf chuck power in the former because of its larger electrode area.

(iv) MESFETs are more sensitive to changes in their channel resistance and transconductance upon plasma expo-

sure than are devices such as heterojunction bipolar transistors (HBTs),²⁹ in which the doping levels are much higher in the individual layers and thus sheet resistance and gain are altered only when comparatively higher amounts of damage have been accumulated. In similar experiments on GaAs/AlGaAs HBTs we observed that changes were minimized at somewhat higher ICP and ECR source powers (~500 W) than seen for MESFETs (~200 W).

ACKNOWLEDGMENTS

The work at UF is partially supported by a DOD University Research Initiative administered through AFOSR (H. C. DeLong, Contract No. F49620-1-0026). The work at SNL is supported by DOE under Contract No. DE-AC04-94AL85000. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed-Martin Company, for the United States Department of Energy. The staff at the Microfabritech facility at UF are gratefully acknowledged.

¹See, for example, M. S. Shur, in *Introduction to Semiconductor Technology, GaAs and Related Compounds*, edited by C. T. Wang (Wiley Interscience, New York, 1990), Chap. 2.

²R. Williams, *Modern GaAs Processing Methods* (Artech House, Boston, 1990).

³Y.-S. Lin and S.-S. Lu, IEEE Trans. Electron Dev. Lett. **EDL-17**, 452 (1996).

⁴S.-J. Chang and C.-P. Lee, Solid-State Electron. **39**, 1015 (1996).

⁵S. M. Lande, *GaAs Electronic Materials and Devices—A Strategic Study of Materials, Technologies and Companies Worldwide* (Elsevier, Oxford, 1996).

⁶K. L. Seaward and N. J. Moll, J. Vac. Sci. Technol. B **10**, 46 (1992).

⁷R. J. Shul, M. E. Sherwin, A. G. Baca, and D. J. Rieger, Electron. Lett. **32**, 70 (1996).

⁸R. M. Olson, Jr., T. E. Kazior, B. Lane, W. M. Holber, and L. Bourget, J. Electrochem. Soc. **143**, 233 (1996).

⁹S. J. Pearton, F. Ren, and C. R. Abernathy, Semicond. Sci. Technol. **8**, 1897 (1993).

¹⁰K. L. Seaward, N. J. Moll, and W. F. Stickley, J. Electron. Mater. **19**, 385 (1990).

¹¹S. W. Pang, J. Electrochem. Soc. **133**, 784 (1986).

¹²C. I. H. Ashby, *Properties of GaAs*, 2nd ed. (EMIS Data Reviews, INSPEC, London, 1990).

¹³J. Asmussen, J. Vac. Sci. Technol. A **7**, 883 (1989).

¹⁴*High Density Plasma Sources*, edited by O. A. Popov (Noyes, Park Ridge, NJ, 1996).

¹⁵F. C. Sze and J. Asmussen, J. Vac. Sci. Technol. A **11**, 1289 (1993).

¹⁶M. A. Leiberlmann and A. Lichtenburg, *Principles of Plasma Discharges and Materials Processing* (Wiley, New York, 1994).

¹⁷S. Gorbalkin, L. Berry, and J. Roberto, J. Vac. Sci. Technol. A **8**, 2893 (1990).

¹⁸J. E. Stevens, J. L. Cecchi, Y. C. Huang, and R. L. Jarecki, J. Vac. Sci. Technol. A **9**, 696 (1991).

¹⁹M. A. Leiberlmann and R. A. Gottscho, *Physics of Thin Films*, edited by M. Francombe and J. Vossen (Academic, New York, 1993), Vol. 17.

²⁰S. Thomas III, K. K. Ko, and S. W. Pang, J. Vac. Sci. Technol. A **13**, 894 (1995).

²¹J. W. Lee, J. Hong, and S. J. Pearton, Appl. Phys. Lett. **68**, 847 (1996).

²²H. P. Gillis, D. A. Choutov, and K. P. Martin, JOM, August 1996, pp. 50–55.

²³S. J. Pearton, Mater. Sci. Rep. **4**, 313 (1990).

²⁴S. K. Ghandhi, *VLSI Fabrication Principles* (Wiley, New York, 1994).

²⁵G. H. Kinchin and R. S. Pease, Rep. Prog. Phys. **18**, 1 (1955).

²⁶See, for example, J. F. Ziegler, J. P. Biersack, and U. Littmark, *The Stopping and Ranges of Ions in Solids* (Pergamon, New York, 1985).

²⁷E. L. Hu, C.-H. Chen, and D. L. Green, J. Vac. Sci. Technol. B **14**, 3632 (1996).

²⁸S. Murad, N. I. Cameron, S. P. Beaumont, and C. D. W. Wilkinson, J. Vac. Sci. Technol. B **14**, 3658 (1996).

²⁹J. W. Lee, C. R. Abernathy, S. J. Pearton, F. Ren, C. Constantine, C. Barratt, and R. J. Shul, Appl. Phys. Lett. **70**, 2410 (1997).

Dry etching of germanium in magnetron enhanced SF₆ plasmas

G. F. McLane,^{a)} M. Dubey, and M. C. Wood
Army Research Laboratory, Fort Monmouth, New Jersey 07703

K. E. Lynch
New Jersey Institute of Technology, Newark, New Jersey 07102

(Received 20 February 1997; accepted 16 May 1997)

Magnetron enhanced reactive ion etching of germanium was investigated in SF₆ plasmas. Ge etch rates were determined as a function of cathode power density (0.1–0.5 W/cm²), pressure (2–8 mTorr), and SF₆ flow rate (2–11.5 sccm). Etch rate increased as pressure and flow rate were increased, but exhibited the unusual characteristic of decreasing as cathode power was increased. Auger electron spectroscopy measurements showed the presence of a sulfur residue (<1 at. %) upon etching, while scanning electron microscopy revealed that smooth etched surfaces were attained in SF₆ magnetron enhanced plasmas. © 1997 American Vacuum Society. [S0734-211X(97)14304-9]

I. INTRODUCTION

Present day electronic computer logic gates with large fanout requirements can have propagation delays up to many hundreds of picoseconds due to electrical capacitance. These delays, independent of individual logic component speeds, severely limit electronic computer performance. Optical interconnects have the capability of dramatically improving electronic computer performance speeds through the reduction of electrical capacitance, along with providing an inherently high bandwidth capability. The digital logic components can be improved in speed and power by direct optical connections at the gate level. With removal of the capacitance in the fanout logic gates and wiring by using optics, the electronic logic can reach its natural switching time of less than 10 ps. In addition, the power dissipation in electronic circuits arises from the charging of these capacitors, and is also reduced with the use of optical interconnects. Very large scale integrated optics provides high bandwidth interconnections using optical waveguides as transmission lines directly connected to the logic gates.¹

The problem of light scattering in the optical waveguides at the wavelengths normally associated with III–V semiconductor light sources can be solved by use of longer wavelength (10 μm) midinfrared light because the scattering losses drop with wavelength to the third power.² Appropriate waveguides can be created with a high refractive index ratio between cladding and guide. A practical implementation of such a waveguide in the midinfrared region uses Ge with refractive index 4.0 on GaAs with refractive index 3.27. Ge has an almost perfect lattice match with GaAs, and can be deposited onto a GaAs substrate via ultrahigh vacuum (UHV) e-beam deposition.³ Subsequent processing of Ge for waveguide formation requires patterning of appropriate structures by dry etching techniques. Since the chlorides and fluorides of Ge are volatile species,⁴ Cl-based and F-based plasmas are obvious choices for use in the dry etching of Ge. Reactive ion etching (RIE) of Ge has been reported in CF₄, SF₆, CCl₂F₂, CF₃Br, and HBr plasmas.^{5–8} This article re-

ports on the magnetron enhanced reactive ion etching (MIE) of Ge in SF₆ plasmas. MIE is similar to conventional RIE but with the presence of a magnetic field close to the wafer surface. This magnetic field confines electrons in closed orbits and increases ionization efficiency, resulting in high density plasmas with high etch rates. MIE has been successfully used in the past to etch silicon⁹ and III–V semiconductors.^{10,11}

II. EXPERIMENT

Magnetron reactive ion etching experiments on germanium were performed in a Materials Research Corp. MIE 710 system described previously.¹² Ge samples having very low absorption in the midinfrared wavelength range were obtained by depositing Ge layers (1 μm thick) onto (100) GaAs substrates at 100 °C using an UHV e-beam deposition system. Patterns were formed on the samples with AZ5214 photoresist, which was removed with an acetone rinse subsequent to the etch. To prevent etching into the substrate, etch durations were chosen to produce etch depths which were less than film thickness. Etch depth was determined from Dektak profilometer measurements. Etch rates were measured as a function of cathode power, pressure, and flow rate in SF₆ plasmas. The Ge layer etch rates were identical to rates obtained on bulk single crystal Ge samples. Auger electron spectroscopy (AES) measurements of etched and control samples were performed in a Perkin Elmer PHI 660 scanning Auger microprobe. Etch profiles were examined by scanning electron microscope measurements.

III. RESULTS

Figure 1 shows Ge etch rate and cathode bias voltage as a function of cathode power density. Germanium fluoride is more volatile than the chlorides or bromides of Ge, and it is reasonable to assume that GeF₄ is the main etch product for Ge etching in SF₆ plasmas.⁷ Increasing cathode power usually results in greater generation of reactant species and higher ion bombardment energies, producing higher etch rates. Bright *et al.*¹³ have reported this type of etch rate dependence on power for RIE of Ge in CF₄/H₂ plasmas. How-

^{a)}Electronic mail: gmlane@arl.mil

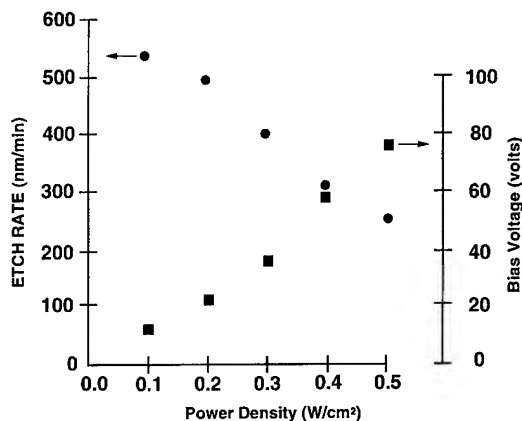


FIG. 1. Germanium etch rate and cathode bias voltage as a function of cathode power density, with 5.0 sccm SF_6 flow rate and 2 mTorr pressure.

ever, Fig. 1 shows that Ge etch rate decreases as power is increased, in spite of higher cathode bias voltages and resultant ion bombardment energies. This type of behavior has been reported previously for RIE of III-V semiconductors in chlorine-based etch gases.¹⁴⁻¹⁶ It was attributed to sputter desorption of the halogen radicals at higher powers and bias voltages, before these radicals could react with the wafer surface.

Thus, under the conditions reported here, Ge etching appears to be adsorption limited above some relatively low cathode bias voltage. Unlike Si etching, Ge etch rates in fluorine-based plasmas are known to depend more on chemical reactivity rather than physical sputtering,⁷ and for MIE of Ge in SF_6 the sputtering component can actually lower the etch rate substantially.

Figure 2 exhibits the flow rate dependence of Ge etch rate. The substantial etch rate increase with flow rate indicates that the etch rate is limited by the SF_6 feed gas supply over the given flow rate range. An etch rate greater than 1 $\mu\text{m}/\text{min}$ was attained when flow rate was increased to 11.5 sccm, with further etch rate increase achievable with even

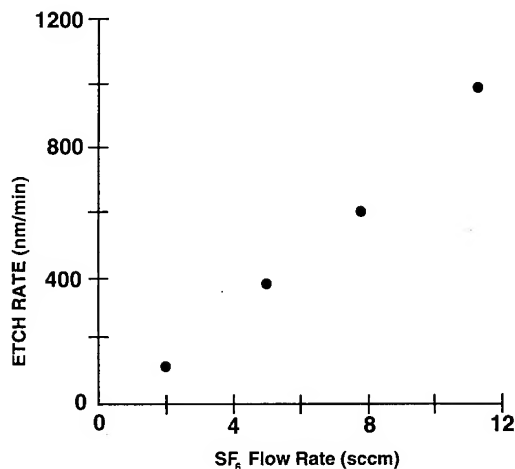


FIG. 2. Germanium etch rate as a function of SF_6 flow rate, with 2 mTorr pressure and 0.3 W/cm^2 cathode power density.

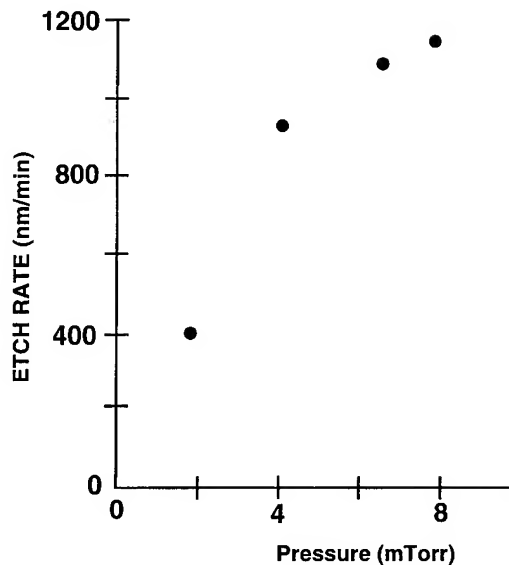


FIG. 3. Germanium etch rate as a function of pressure, with 5.0 sccm SF_6 flow rate and 0.3 W/cm^2 cathode power density.

higher flow rates. The dependence of etch rate with chamber pressure is shown in Fig. 3. Etch rate increases significantly as pressure is raised, reaching 1.2 $\mu\text{m}/\text{min}$ at about 8 mTorr, whereupon it shows an indication of approaching saturation. Etch rate increase with pressure is due to an increase in reactant fluorine species production resulting from greater availability of the SF_6 source gas. The results here show that etch rates above 1 $\mu\text{m}/\text{min}$ are easily achieved with MIE SF_6 plasmas. In comparison, Oehrlein *et al.*⁷ obtained a Ge etch rate of 0.6 $\mu\text{m}/\text{min}$ using RIE with SF_6 at 150 mTorr, 150 W, 100 sccm etch conditions.

AES surface measurements performed on a Ge sample etched in SF_6 revealed the presence of a sulfur residue with a concentration of <1 at. %, but no evidence of the presence of fluorine. Involatile germanium sulfides are known to inhibit Ge etching in SF_6/H_2 plasmas,⁷ but should have no

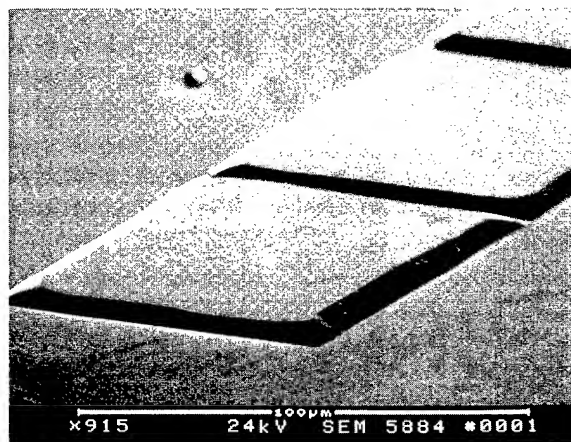


FIG. 4. Scanning electron micrograph of Ge sample patterned with AZ5214 photoresist and etched in SF_6 with 5.0 sccm flow rate, 2 mTorr pressure, and 0.3 W/cm^2 cathode power density. Etch depth is 2.75 μm .

effect at the low sulfur concentrations found here. Figure 4 shows a scanning electron micrograph of a Ge sample patterned with AZ5214 photoresist and etched in SF₆. The etched surface is smooth, with a negative undercut profile caused mostly by photoresist edge erosion.

IV. SUMMARY

In summary, magnetron enhanced SF₆ plasmas have produced high Ge etch rates with smooth surfaces. The Ge etch mechanism appears to be adsorption limited under the conditions studied, with Ge etch rate decreasing as cathode power and bias voltage is increased.

¹L. C. West, C. W. Roberts, E. C. Piscani, M. Dubey, K. A. Jones, and G. F. McLane, SPIE International Symposium OE/LASE, San Jose, CA, 27 January-2 February 1996 (unpublished).

²D. Marcuse, Bell Syst. Tech. J. **48**, 3233 (1969).

³M. Dubey, G. F. McLane, K. A. Jones, L. C. West, C. W. Roberts, and E. C. Piscani, SPIE International Symposium OE/LASE, San Jose, CA, 27 January-2 February 1996 (unpublished).

⁴*Handbook of Chemistry and Physics*, 67th ed., edited by R. C. Weast (Chemical Rubber, Boca Raton, FL, 1987).

⁵G. S. Oehrlein, T. D. Bestwick, P. L. Jones, and J. W. Corbett, Appl. Phys. Lett. **56**, 1436 (1990).

⁶G. S. Oehrlein, Y. Zhang, G. M. W. Kroesen, E. de Fresart, and T. D. Bestwick, Appl. Phys. Lett. **58**, 2252 (1991).

⁷G. S. Oehrlein, T. D. Bestwick, P. L. Jones, M. A. Jaso, and J. L. Lindstrom, J. Electrochem. Soc. **138**, 1443 (1991).

⁸Y. Zhang, G. S. Oehrlein, and E. de Fresart, J. Appl. Phys. **71**, 1936 (1992).

⁹A. A. Bright, S. Kausik, and G. S. Oehrlein, J. Appl. Phys. **62**, 2518 (1987).

¹⁰G. F. McLane, M. Meyyappan, H. S. Lee, M. W. Cole, D. W. Eckart, R. T. Lareau, M. Namaroff, and J. Sasserath, J. Vac. Sci. Technol. A **11**, 333 (1993).

¹¹G. F. McLane, L. Casas, S. J. Pearton, and C. R. Abernathy, Appl. Phys. Lett. **66**, 3328 (1995).

¹²G. F. McLane, M. W. Cole, D. W. Eckart, P. Cooke, R. Moerkirk, and M. Meyyappan, J. Vac. Sci. Technol. B **11**, 1753 (1993).

¹³A. A. Bright, S. S. Iyer, S. W. Robey, and S. L. Delage, Appl. Phys. Lett. **53**, 2328 (1988).

¹⁴S. J. Pearton, U. K. Chakrabarti, W. S. Hobson, and A. P. Kinsella, J. Vac. Sci. Technol. B **8**, 607 (1980).

¹⁵Y. Z. Juang, Y. K. Su, S. C. Shei, and B. C. Fang, J. Vac. Sci. Technol. A **12**, 75 (1994).

¹⁶R. J. Contolini and A. J. D'Ásaro, J. Vac. Sci. Technol. B **4**, 706 (1986).

Spatial variation of the etch rate for deep etching of silicon by reactive ion etching

Bo Asp Møller Andersen,^{a)} Ole Hansen, and Martin Kristensen
Mikroelektronik Centret, Technical University of Denmark, 2800 Lyngby, Denmark

(Received 8 March 1996; accepted 30 May 1997)

The macroscopic uniformity of deep etching into silicon by reactive ion etching (RIE) with a $\text{SF}_6\text{--O}_2$ plasma was studied. The spatial variation of the etch rate across a 4 inch wafer in a single wafer system is a function of the process parameters and the configuration of the etch chamber. It was found that, for a constant load of silicon exposed to the plasma, the etch rate variation can be controlled through the applied rf power, the chamber pressure, and the gas mixture. It was also found that the etch rate uniformity varies with the load of silicon exposed to the plasma. The result is a balance between the flux of neutral radicals and the flux of energetic ions to the surface. This balance is due to the RIE etch mechanism, which involves synergism between the two fluxes.

© 1997 American Vacuum Society. [S0734-211X(97)15204-0]

I. INTRODUCTION

Reactive ion etching (RIE) is an important process for the fabrication of many integrated semiconductor devices. RIE is a plasma-assisted process, that involves interaction of a glow discharge with a solid substrate. The plasma serves as a source for reactive species by dissociation and ionization of the gases. The neutral radicals react with the solid to form volatile products, which are pumped away. The ionization in the plasma gives rise to a bombardment of the substrate due to the self-bias voltage across the plasma sheath. The ion bombardment enhances the etch rate and the anisotropy of the process. Coburn and Winters¹ found that an etch rate with a simultaneous flux of ions and radicals is larger than the sum of the etch rates from individual fluxes.

The main parameters of the process are the gas mixture, the applied rf power, and the chamber pressure. Through these parameters the amount of radicals and the ion bombardment can be controlled. Other parameters such as the configuration of the etch chamber, the masking material on the wafer, and the design of the mask pattern will also affect the etch process. These numerous interrelated parameters in RIE complicate general use of the process. In addition, different properties of the etch process have to be optimized. These include the selectivity between the mask material and the material to be etched, the sidewall profile of the etched structures, the etch rate, and the uniformity.

For many applications a precise etched depth is most important. Methods have been developed to perform *in situ* measurements of the etch depth. These techniques only measure the etch depth at a single point or at a few selected points. The resulting accuracy is therefore limited by the uniformity of the RIE process.

Different types of uniformity in the RIE process have to be taken into account. These can roughly be divided into macrouniformity and microuniformity. The microuniformity is a feature size dependent parameter that is given by the variation in etch depth as a function of the linewidth. The

macrouniformity is the overall uniformity of identical structures as a function of position on the wafer.

Fluorine chemistry has been widely used for RIE etching of silicon, and gas mixtures of SF_6 and O_2 have been especially popular. Other fluorine gases such as CF_4 , CHF_3 and C_2F_6 have also been widely used. In this article the spatial variation of etch rates was investigated for RIE of silicon in a $\text{SF}_6\text{--O}_2$ plasma. The variation was studied as a function of the ratio between SF_6 and O_2 in the input gas mixture, the applied rf power, the chamber pressure, and the loading effect of exposed silicon area.

II. EXPERIMENT

A. Equipment

The experiments were carried out in a Surface Technology System cluster system consisting of a RIE chamber and a plasma enhanced chemical vapor deposition (PECVD) chamber connected to a loadlock through a closed transfer chamber. The RIE chamber has parallel plate electrodes connected to a 13.56 MHz rf generator through an automatic matching network. The chamber is made of aluminum and is electrically connected to the grounded aluminum top electrode. The aluminum bottom electrode has a diameter of 27 cm. A 1.0 cm thick shield plate made of quartz with a 4 inch hole in the center covers the bottom electrode. The electrode spacing is 9 cm. An oil bath system cools the bottom electrode to 30 °C. The gases are let into the chamber through the top electrode which has a showerhead. The flow of each gas and the chamber pressure can be controlled independently.

B. Wafer preparation

The experiments were performed with lightly doped 4 inch silicon wafers. The masking material was a thermal oxide. The wafers were cleaned in a standard RCA process and thermally oxidized to a thickness of 1 μm . After the thermal process the wafers were treated with hexamethyldisilazane (HMDS) and 1.4 μm of photoresist was spun onto the wa-

^{a)}Electronic mail: bam@mic.dtu.dk

fers. The resist was patterned through a standard photolithographic process. The mask pattern consisted of an array of 7 by 7 mm chips covering the entire wafer. The resist pattern was transferred to the oxide by wet etching in buffered hydrofluoric acid, and then stripped off before the silicon etch process.

C. RIE chamber conditioning

Cleaning and preconditioning of the etch chamber was performed before each experiment. The chamber was cleaned by an oxygen plasma for 10 min followed by a 5 min preconditioning with the process used in the experiment. Both the cleaning and the preconditioning were carried out with a blank 4 inch silicon wafer in the chamber.

D. Measurements

The oxide mask was removed by immersing in hydrofluoric acid after the RIE etch. Etch depths were measured at identical structures in 14 positions on each wafer with a TenCor stylus step profiler.

E. Process parameters

The RIE etch time was 60 min in all processes. The etch process was investigated by varying the applied rf power, chamber pressure, oxygen flow, and the surface area of exposed silicon. The chamber configuration was unchanged throughout the complete series of experiments.

III. RESULTS

Through preliminary etch experiments a process was developed that yields good uniformity. This process was used as a starting point for investigating the impact of the different parameters. The variation of etch depths across the wafers are now discussed.

A. Power variation

The influence of the applied rf power on the etch rate was investigated using a mask with a surface coverage of 50%. The gas mixture was 50 sccm SF_6 and 20 sccm O_2 at a pressure of 150 mTorr. Rf power levels of 32 W, 40 W, 50 W, and 60 W were used. In Figure 1 the etch depths measured are shown as a function of radial position with the applied rf power as the parameter.

At high rf power (50 W and 60 W) the etch rate is lowest at the center of the wafer, and increases monotonically towards the edge. At an applied rf power of 40 W the etch rate is most uniform. The variation of the etch depth is less than 1% within a diameter of 8 cm. The etch rate decreases close to the edge. At 32 W the etch rate is highest at the center of the wafer and it monotonically decreases towards the edge.

B. Pressure variation

The influence of the chamber pressure on the etch rate was investigated using a mask with a surface coverage of 50%. The applied rf power and the gas mixtures were 40 W, 50 sccm SF_6 , and 20 sccm O_2 , respectively. Experiments

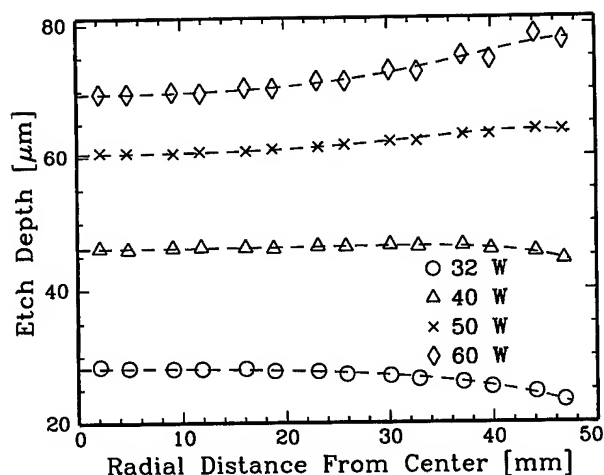


FIG. 1. Radial variation of the etch depth with the applied rf power as the parameter. The measured etch depths are shown by symbols. The dashed lines are fits to the etch depths according to Equation (14).

were performed with chamber pressures of 100 mTorr, 120 mTorr, 150 mTorr, and 165 mTorr. The etch depths measured are shown in Figure 2 as a function of radial position with the chamber pressure as the parameter.

At a pressure of 100 mTorr the etch rate is lowest at the center of the wafer and has monotonically increasing etch rate towards the edge. At 120 mTorr and 150 mTorr the etch rate is again lowest in the center. The etch rate increases towards the edge of the wafer, it passes through a maximum, and then decreases at the edge. When the pressure is increased further to 165 mTorr, the etch rate is highest in the center, and decreases monotonically towards the edge of the etched wafer.

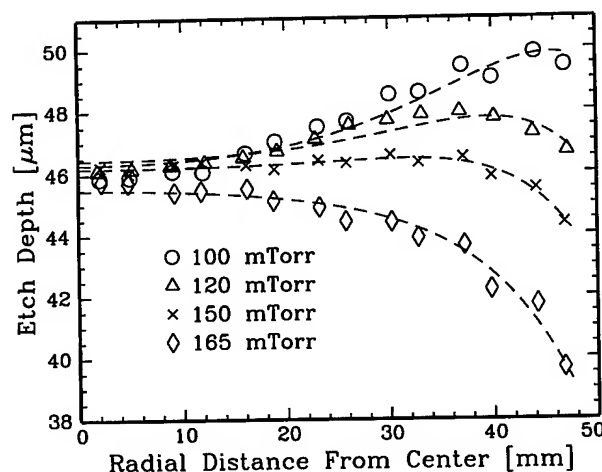


FIG. 2. Radial variation of the etch depth with the chamber pressure as the parameter. The measured etch depths are shown by symbols. The dashed lines are fits to the etch depths according to Equation (14).

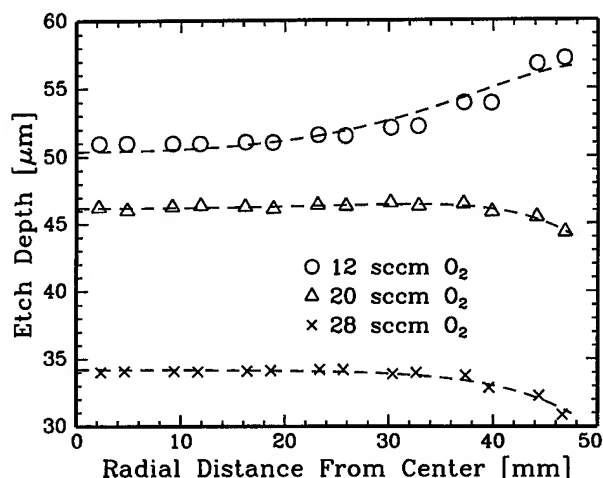


FIG. 3. Radial variation of the etch depth with the oxygen flow as the parameter. The measured etch depths are shown by symbols. The dashed lines are fits to the etch depths according to Equation (14).

C. Variation of the oxygen content

The influence of the gas mixture on the etch rate was investigated using a mask with a surface coverage of 50%. The applied rf power and the chamber pressure were 40 W and 150 mTorr, respectively. The SF_6 flow was kept constant at 50 sccm. The etch rate variation was investigated for oxygen flows of 12 sccm, 20 sccm, and 28 sccm. The etch depths measured are shown in Figure 3 as a function of radial distance with the oxygen flow as the parameter.

At a low oxygen flow rate (12 sccm) the etch rate is lowest at the center of the wafer, with a monotonic increase towards the edge. When the oxygen flow is increased to 20 sccm the variation of the etch rates shifts. There is a local minimum of the etch rate in the center of the wafer. The etch rate increases towards the edge, goes through a maximum, and then decreases close to the edge. With the oxygen flow increased to 28 sccm a similar spatial variation is observed.

D. Variation of the exposed silicon area

The etch process was investigated for different areas of silicon exposed to the plasma. The remaining plasma parameters were fixed at 40 W of applied rf power, a chamber pressure of 150 mTorr, and gas flows of 50 sccm SF_6 and 20 sccm O_2 . In Figure 4 the etch depths measured are shown as function of radial distance with the area of exposed silicon as the parameter. Three masks which exposed 30%, 50%, and 70% of the wafer area were used.

For an exposed area of 30% the etch rate is highest at the center of the wafer and it decreases monotonically towards the edge. At 50% load the etch rate is reasonably uniform within a central diameter of 8 cm. The variation of the etch depth is only 1% within this area. The etch rate falls off close to the edge of the wafer. At an exposed area of 70% the etch rate increases monotonically from the center towards the edge.

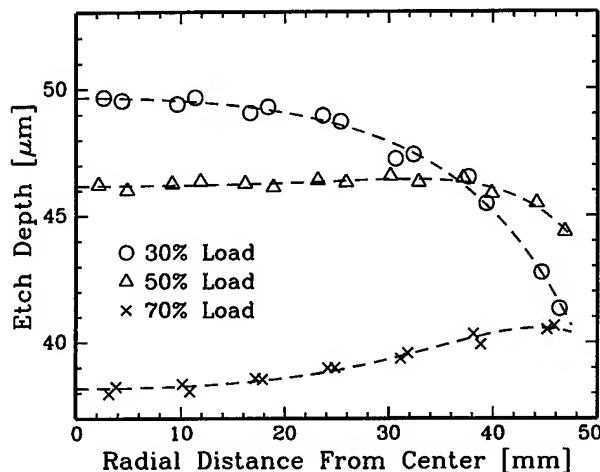


FIG. 4. Radial variation of the etch depth with the percentage of the wafer exposed as the parameter. The measured etch depths are shown by symbols. The dashed lines are fits to the etch depths according to Equation (14).

IV. DISCUSSION

The spatial variation of the etch depths in Figures 1–4 depends systematically on the process parameters. Changing the process parameters will change the plasma in the etch reactor and hence the characteristics of the etch process. In the following we will develop a simple model for a connection between process parameters, changes in the plasma, and spatial variation in the etch depth. With this simple model we show that the etch rate variation across a wafer can be explained qualitatively.

The surface processes involved in plasma-assisted etching have been discussed by several researchers. Reactive ion etching can be driven in three different regimes characterized by physical sputter etching, ion assisted chemical etching, and pure chemical etching as the dominant etching mechanisms. Winters *et al.*² have reviewed the etch mechanisms and provided evidence that RIE relies on an enhancement of the chemical etching by the energy transferred to the surface by bombarding ions. The bombarding ions induce a chemical reaction, which produces a weakly bound molecule that desorbs from the surface of the substrate. This process is called chemical sputtering.

Reactive ion etching is most efficient in the chemical sputtering regime and most RIE processes are therefore driven in this regime. In the chemical sputtering regime the etch rate can be expressed as a combination of the volume of silicon removed per incident ion and the volume of silicon removed per neutral radical sticking to the surface. The etch rate will then be a function of the ion energy, the ion flux, and the neutral flux. Gottscho *et al.*³ have given a simple expression where the etch rate is expressed as

$$R = \frac{\kappa E_i J_i}{1 + \kappa E_i J_i / v S_0 J_n}, \quad (1)$$

where κ is the volume of silicon removed per unit bombardment energy and ion, E_i is the ion energy, J_i is the ion flux to the surface, v is the volume removed per reacting neutral,

S_0 is the sticking probability of neutrals to the silicon surface, and J_n is the flux of neutrals to the surface.

The etch rate will become very small in the two other regimes where one etching mechanism, either the chemical etching by neutrals or the sputter etching by energetic ions, becomes dominant.

The measured variation of the etch rate across the etched wafers can be explained through Equation (1), if the spatial variation of the contribution from the neutrals and the ions to the etch rate can be found. The flux of neutral radicals to the surface, and hence the contribution to the etch rate, is assumed to be directly proportional to the neutral concentration. The spatial variation of the ion bombardment across the etched wafer can originate from concentration differences and from distortion of the electrical field due to the configuration of the etch chamber.

A model describing the fluorine concentration in the chamber can be obtained from a steady state solution to the continuity equation including generation, loss, and diffusion of fluorine. If recombination of fluorine in the bulk of the plasma is ignored, the steady state fluorine concentration C is determined by the continuity equation,

$$\frac{\partial C}{\partial t} = D \cdot \nabla^2 C + G = 0, \quad (2)$$

where G is the volume generation rate of fluorine in the plasma, and D is the diffusivity of fluorine in the gas phase. The etching reactor has circular symmetry. Equation (2) is therefore most conveniently solved using cylindrical coordinates, where the angular dependence can be ignored. Equation (2) can be solved for C with the appropriate boundary conditions at the center of the chamber ($r=0$), at the pump ($r=R_2$), and at the wafer and electrode surfaces. In the center of the chamber there is no lateral flux,

$$\frac{\partial C(0, \Theta, z)}{\partial r} = 0. \quad (3)$$

At the lower horizontal surface the drain of fluorine is given by the reaction coefficient and the fluorine concentration,

$$D \frac{\partial C(r, \Theta, 0)}{\partial z} = k(r) C(r, \Theta, 0), \quad (4)$$

where $k(r)$ is the fluorine reaction rate at the wafer and at the shield surface. Surface reactions at the top electrode are neglected, hence

$$D \frac{\partial C(r, \Theta, h)}{\partial z} = 0, \quad (5)$$

where h is the electrode spacing. A boundary condition has to be selected for the outer rim of the bottom electrode ($r=R_2$). An assumption could be made that the pump is totally absorbing at the edge, which will take the concentration close to zero. This assumption, however, requires a pumping speed three orders of magnitude higher than the actual pumping speed. Hence, it is more appropriate to assume that the pump is nonabsorbing, which will give a zero radial flux at the edge and the boundary condition at $r=R_2$ will be

$$\frac{\partial C(R_2, \Theta, z)}{\partial r} = 0. \quad (6)$$

In reality there will be some loss due to recombination on surfaces which is not considered in this model. However, this will not influence the form of the spatial fluorine concentration distribution significantly in the vicinity of the wafer placed in the center of the reactor.

Even with these simple boundary conditions Equation (2) is difficult to solve for the cylindrical geometry of the etching chamber. The main features of the solution are preserved, however, even if rather crude approximations are made. If the average fluorine concentration in the z direction is obtained from integration of Equation (2), the loss of fluorine to the surface will appear as a bulk recombination term,

$$\frac{k(r)C(r)}{h} = \frac{C(r)}{\tau(r)}. \quad (7)$$

Then Equation (2) can be transformed into a one dimensional equation (for convenience the z average of C is not explicitly stated):

$$\frac{\partial C}{\partial t} = D \nabla_r^2 C + G - \frac{1}{\tau} C = 0. \quad (8)$$

The stationary solution can be found with the boundary conditions,

$$\frac{\partial C(0)}{\partial r} = 0, \quad \frac{\partial C(R_2)}{\partial r} = 0, \quad (9)$$

$$k(r < R_1) = k_1, k(R_1 < r < R_2) = k_2,$$

where R_1 is the wafer radius (50 mm). Furthermore C and $\partial C/\partial r$ have to be continuous as function of the radius. Equation (8) can be simplified into

$$D \left\{ \frac{\partial^2 C}{\partial r^2} + \frac{1}{r} \frac{\partial C}{\partial r} \right\} - \frac{1}{\tau} C = -G. \quad (10)$$

Equation (10) has to be solved for the two different areas of the bottom electrode, which are over the etched wafer and over the quartz shield, respectively. The solution to the homogeneous equation is a sum of modified Bessel functions. The general solutions (11) and (12) give the fluorine concentration variation over the entire bottom electrode:

for $0 \leq r \leq R_1$

$$C_1(r) = G \left\{ A_1 \cdot I_0 \left(\frac{r}{\sqrt{D\tau_1}} \right) + \tau_1 \right\}, \quad (11)$$

and

for $R_1 \leq r \leq R_2$

$$C_2(r) = G \left\{ A_2 \cdot I_0 \left(\frac{r}{\sqrt{D\tau_2}} \right) - B_2 \cdot K_0 \left(\frac{r}{\sqrt{D\tau_2}} \right) + \tau_2 \right\}. \quad (12)$$

The coefficients, A_1 , A_2 , and B_2 , are functions of τ_1 , τ_2 , R_1 , R_2 . The fluorine concentration given by Equations (11) and (12) is shown in Figure 5 where τ_2 was set to $\tau_2 = \tau_1/50$, which corresponds to the selectivity in the etch

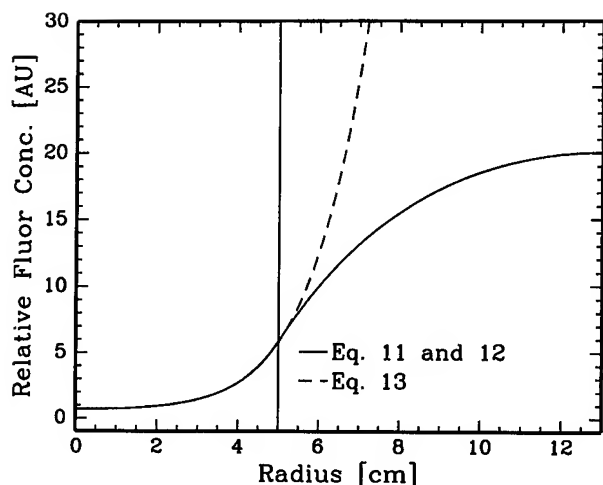


FIG. 5. The fluorine concentration in the etch reactor as calculated by Equations (11) and (12) as a function of the radius from the center (solid curve). The solid vertical line marks the edge of the wafer. The dashed curve shows the approximated fluorine concentration as calculated by Equation (13).

process between silicon and silicon dioxide. It can be observed that the concentration of neutral radicals above the shield is higher than above the wafer. Above the wafer the concentration is highest at the edge of the wafer, with a monotonic decrease towards the center. This corresponds well with the concentration variation of neutral radicals found by numerical calculations for a plasma etch system with a similar shield configuration by Kobayashi *et al.*⁴

For use in the etch rate model, Equation (11) which covers the area of the wafer, can be approximated by an expansion into a series given by

$$C(r) = G \left\{ k_1 + a \cdot r^2 + \frac{a^2}{16} \cdot r^4 + \frac{a^3}{576} \cdot r^6 \right\}, \quad (13)$$

where the coefficients k_1 and a are functions of τ_1 and D . This approximation to $C(r)$ is also shown in Figure 5. The two curves coincide for $r < R_1$.

The contribution from the ion bombardment to the spatial variation of the etch rate across the wafer originates from the chamber configuration. The 10 mm thick quartz shield plate placed on the bottom electrode gives rise to distortion of the electrical field. This was confirmed by visual inspection of the dark space of the plasma through a window in the side-

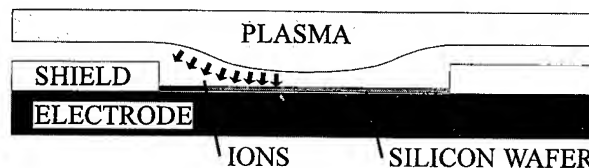


FIG. 6. A schematic view of a cross section of the RIE chamber. The deflection of the glow region of the plasma that is observed will result in a variation of the angular incidence of the ions.

wall of the chamber. The plasma distortion is shown schematically in Figure 6. The thickness of the dark space is of the order of 10–15 mm over the range of parameters investigated in this work.

The distortion of the plasma will result in an angular incidence of the ions, with an increasing angle from normal incidence on the wafer towards the edge. Mayer and Barker⁵ have found that the etch efficiency of the ion bombardment in fluorine based RIE systems decreases monotonically with an increasing angle of incidence. The spatial variation of the etch rate due to the ion bombardment will enter Equation (1) through the product $\kappa E_i J_i$.

Matsuo⁶ observed this effect in an investigation of RIE etching of SiO_2 . The SiO_2 was etched by the reactive ion CF_n^+ and the contribution to the etch process from neutrals was insignificant. Variation of the etch rate induced by a 0.5 mm high step was entirely due to variations in the ion bombardment.

It is difficult to measure the spatial variation of the ion contribution to the etch rate. Instead, the variation can be estimated from the measured etch depths for a process where the ion bombardment is the etch rate limiting factor. This is the case in the process where the chamber pressure is increased to 165 mTorr. There is an excess of fluorine atoms across the entire wafer and, simultaneously, the ion energy decreases so that the ion bombardment is the limiting factor. A function approximating the variation of the etch rate in this case, and thus the product $\kappa J_i E_i$, can be found by fitting a function to the measured etch depths. A fourth order polynomial of the type $c_0(1 + c_1 \cdot r^2 + c_2 \cdot r^4)$, where r is the distance from the center, is a good fit.

Equation (1) can now be rewritten to give the etch rate variation as function of the radius from the center of the etch chamber.

$$R(r) = \frac{c_0(1 + c_1 \cdot r^2 + c_2 \cdot r^4)}{1 + c_0(1 + c_1 \cdot r^2 + c_2 \cdot r^4) / [k_0 \cdot (k_1 + a \cdot r^2 + a^2/16 \cdot r^4 + a^3/576 \cdot r^6)]}, \quad (14)$$

where k_0 represents $v S_0 G$ from Equation (1). The coefficients in Equation (14) ($c_0, c_1, c_2, k_0, k_1, a$) are difficult to determine directly from the reactor geometry and the process parameters. Instead, these coefficients can be obtained by using Equation (14) to fit the center process used in the in-

vestigations. The coefficients obtained from a fit to the central process are listed in Table I.

Curves describing the spatial variation of the etch rate are obtained by fitting Equation (14) to the experimental data. The fits are obtained with c_0 and k_0 as the only free param-

TABLE I. Values obtained by fitting Equation (14) to the measured etch depth for the central process.

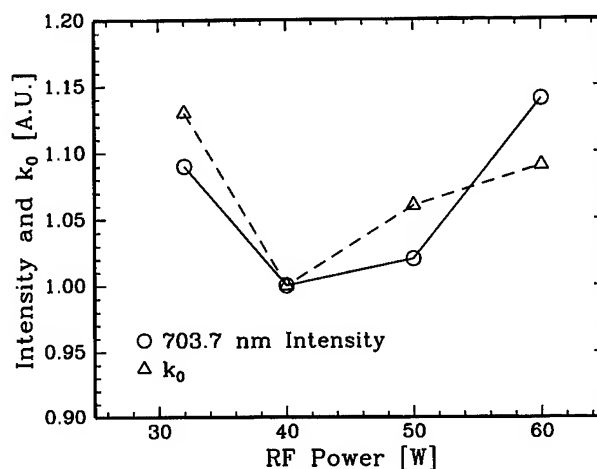
$c_{0,c}$ ($\mu\text{m/h}$)	$c_{1,c}$ (cm^{-2})	$c_{2,c}$ (cm^{-4})	$k_{0,c}$ ($\mu\text{m/h}$)	$k_{1,c}$	a_c (cm^{-2})
103.3	-5.58E-4	-5.91E-4	84.5	.99	3.26E-3

eters, whereas the shape parameters c_1 , c_2 , k_1 , and a are kept at the values obtained from the center process. These curves are shown as dashed curves in Figures 1–4. We find that the curves obtained from the model [Equation (14)] coincide well with the experimental data. The obtained values for c_0 and k_0 relative to the center values are shown in Table II.

From Table II it can be seen that the value of c_0 describing the ion bombardment increases with increasing rf power. This is due to the increase in the ion energy, which is the main change in the plasma, as the dc self-bias increases with increasing rf power.⁷ The value of k_0 describing the change in the contribution to the etch rate from the neutrals is almost independent of the rf power. k_0 is at a maximum for 32 W, decreases slightly as the power is increased to 40 W, and then increases slightly from there as the power is increased further to 50 W and 60 W. This coincides with the measured intensity of the 703.7 nm emission line during the etch processes. This intensity is related to the concentration of the atomic fluorine in the chamber. The value of k_0 and the intensity of the 703.7 nm line are shown in Figure 7 as function of applied rf power. The relatively high fluorine concentration for 32 W applied power is assumed to be due to a low consumption of fluorine due to oxygen passivation of the silicon surface. As the ion energy is low, the removal of the passivation layer is inefficient.

For variation in the chamber pressure c_0 and k_0 vary monotonically with the pressure. With increasing pressure the dc self-bias decreases due to the decrease in the mean free path, which will decrease the energy of the bombarding ions, and hence cause c_0 to decrease. Simultaneously the generation of fluorine atoms will increase with increasing pressure,⁸ which will cause k_0 to increase.

There are three effects of adding oxygen to the gas mixture. One effect in the plasma is a decrease in the ion density,⁷ which results in a decrease of c_0 with increasing oxygen content. A second effect of adding oxygen is an initial increase in the concentration of free atomic fluorine. The

FIG. 7. The measured emission signal from atomic fluorine and k_0 relative to the values for the center process.

oxygen will react with unsaturated compounds of SF_x and prevent recombinations with atomic fluorine. The result will be an initial increase of k_0 with the addition of oxygen. The third effect of adding oxygen is a passivation of the silicon surface due to adsorption of oxygen atoms. This results in a competition between fluorine and oxygen for surface sites and hence a decrease in k_0 for high flows of added oxygen.

When the area of exposed silicon is increased and the other process parameters are kept fixed, the consumption of fluorine atoms will increase. Hence, with increasing silicon area the fluorine concentration decreases, leading to a decrease of k_0 . An increase of the silicon area results in less oxygen passivation, leading to an increased etch efficiency by the bombarding ions, and hence an increase of c_0 .

Good predictions for the etch rate uniformity using other process parameters can be obtained from Equation (14) by using coefficients obtained from an interpolation between the extracted coefficients. The coefficients will depend strongly on the configuration of the etch chamber. If the shield around the etched wafer has a low consumption of fluorine, as is assumed here, the high fluorine concentration at the edge of the wafer has to be compensated by a low ion bombardment efficiency at the edge. If, on the other hand, the shield should consume an amount of fluorine comparable to the wafer,

TABLE II. Values obtained by fitting Equation (14) to the measured etch depths relative to the central process.

	rf power (W)				Silicon load (%)		
	32	40	50	60	30	50	70
$c_0/c_{0,c}$	0.386	1	1.86	2.89	0.69	1	1.21
$k_0/k_{0,c}$	1.13	1	1.06	1.09	1.91	1	0.66
	Pressure (mTorr)				Oxygen flow (sccm)		
	100	120	150	165	12	20	28
$c_0/c_{0,c}$	1.56	1.20	1	0.73	2.17	1	0.62
$k_0/k_{0,c}$	0.79	0.89	1	1.37	0.79	1	0.88

then a flat shield would be preferable in order to give a uniform ion bombardment efficiency. It is possible to find expressions for the parameters in Equation (14) for any given etch chamber geometry.

V. CONCLUSIONS

In this work the relation between the process parameters and the spatial variation of etch rates across the etched wafer was investigated for RIE of silicon by SF_6 - O_2 gas mixtures in a single wafer system. It was found that the variation of the etch rate across the wafer depends strongly on the process parameters and the configuration of the chamber. The RIE process is a synergism between the neutral radicals and the bombarding ions in the chamber. It was shown that both the flux of atomic fluorine to the surface and the etch efficiency of the incident ions have spatial variations. Models for calculations of etch rates have to incorporate the spatial variation, by accounting for the chamber configuration, the mask design on the etched wafer, and the balance between the neutral radicals and the ions as given by the process parameters.

The material of the shield surrounding the wafer on the bottom electrode has a large influence on concentration variation of the neutral radicals in the gas phase above the wafer. The geometric shape of the shield affects the ion bombardment, and induces a spatial variation of the etch efficiency of the incident ions. In the circular symmetric etch chamber used for these investigations the spatial variations are found to be center symmetric. Through variations of the rf power, the pressure, the gas mixture, and the load of silicon exposed to the plasma it was possible to shift the etch rate from having a maximum at the center of the wafer to a maximum at the edge of the wafer.

¹J. W. Coburn and H. F. Winters, *J. Vac. Sci. Technol.* **16**, 391 (1979).

²H. F. Winters, J. W. Coburn, and T. J. Chuang, *J. Vac. Sci. Technol. B* **1**, 469 (1983).

³R. A. Gottscho, C. W. Jurgensen, and D. J. Vitkavage, *J. Vac. Sci. Technol. B* **10**, 2133 (1992).

⁴J. Kobayashi, N. Nakazato, and K. Hiratsuka, *J. Electrochem. Soc.* **136**, 1781 (1989).

⁵T. M. Mayer and R. A. Barker, *J. Vac. Sci. Technol.* **21**, 757 (1982).

⁶S. Matsuo, *J. Vac. Sci. Technol.* **17**, 587 (1980).

⁷P. M. Kopalidis and J. Jorne, *J. Electrochem. Soc.* **139**, 839 (1992).

⁸B. E. Thompson and H. H. Sawin, *J. Electrochem. Soc.* **133**, 1887 (1986).

Efficiency evaluation of postetch metal stack anticorrosion treatments using chemical analyses by x-ray photoelectron spectroscopy and wide dispersive x-ray fluorescence

P. Czuprynski^{a)} and O. Joubert

GRESSI-CNET, France Telecom, BP 98, 38243 Meylan cedex, France

M. Heitzmann, D. Louis, C. Vizios, and E. Lajoinie

GRESSI-LETI (CEA Grenoble), 38054 Grenoble cedex 09, France

(Received 20 March 1997; accepted 23 May 1997)

Chemical analyses by x-ray photoelectron spectroscopy (XPS) combined with wide dispersive x-ray fluorescence (WDXRF) have been used to evaluate the efficiency of metal stack anticorrosion treatments previously etched in a low-pressure high-density plasma source. Analyses demonstrate that residual chlorine is still present in the metal stack showing that corrosion could occur in the subsequent technological steps of the process. In particular, XPS analyses have shown that, after anticorrosion treatment, residual chlorine species are located on the aluminum sidewalls of the features. Combining the results obtained by WDXRF and XPS has allowed a better understanding of the action of anticorrosion treatments. © 1997 American Vacuum Society.

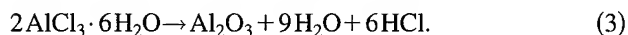
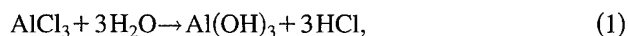
[S0734-211X(97)14604-2]

I. INTRODUCTION

After dry etching, when metal stack interconnects are exposed to atmospheric moisture, rapid corrosion of the etched structures may occur, leading to integrated circuit failure and yield degradation. Furthermore, the continuous reduction in device dimensions imposes at the same time an increase in the number and complexity of metal stacks for interconnections. To address all corrosion related issues becomes critical.¹⁻⁴

The corrosion of patterned aluminum line interconnects occurs when chloride compounds present on the wafer surface after etching hydrolyze by reaction with ambient moisture after being removed from the etching system.⁵ The hydrolysis reaction generates hydrochloric acid that can subsequently react with the aluminum sidewalls.

The typical reaction scheme usually proposed is the following:⁵



The hydrochloric acid resulting from this reaction path is able to dissolve aluminum from the metal stacks after the etching process. Another important point resulting from this reaction path is that it is self-sustained. Once started, the corrosion of aluminum continues as long as moisture is available.

Assessment of the degree of corrosion on etched wafers often relies on scanning electron microscopy views of corrosion sites on conductive lines. However, even if the corrosion reaction is not visible during an inspection, corrosion may still occur after encapsulation of the circuit. Further-

more, metallization schemes are becoming increasingly complex: copper is being alloyed with aluminum, and layers of titanium nitride (TiN) and titanium are being used as antireflective coatings or diffusion barriers. Such metallization schemes have a greater susceptibility to corrosion: chlorides of copper are nonvolatile and may remain on the surface of the wafer after etching, or galvanic cell formation can occur when dissimilar metals are present. The subjective nature of corrosion has led to the development of practical techniques that enable a more reliable evaluation of the effectiveness of postetch corrosion treatments such as ion chromatography,³ thermal desorption spectroscopic analysis,¹ x-ray fluorescence techniques, or x-ray photoelectron spectroscopy techniques.⁷

Since the presence of chlorine on metal surfaces can lead to corrosion, postetch anticorrosion treatments have to be efficient at removing all chlorine-based compounds from all surfaces of the metal stacks. As the anisotropy of the etching is obtained by forming chlorine-rich carbon films on the aluminum and resist sidewalls of the features, anticorrosion treatments must decrease the chlorine concentration as much as possible throughout the thickness of the sidewall passivation polymer formed.^{6,7} Modern anticorrosion treatments consist in many cases of high-pressure H₂O/O₂-based plasmas followed by a depolymerizing wet chemistry. The role of the water-based plasma treatment is to strip the photoresist (which after being exposed to chlorine-rich plasmas contains a large amount of chlorine) and lower the chlorine content on surfaces by forming volatile HCl. The role of the wet depolymerizing chemistry is to remove the chlorine-rich aluminum species left on the metal surfaces after anticorrosion plasma treatments. After this process, no chlorine should remain on any metal surface of the stack.

In this work, metal stacks were etched in a commercial inductively coupled plasma reactor (ICP) using the widely

^{a)}Electronic mail: pascal.czuprynski@cnet.francetelecom.fr

used Cl_2/BCl_3 chemistries.^{8,9} After etching, anticorrosion treatments were performed *in situ* in H_2O - and O_2 -based plasmas followed by a water rinse and depolymerizing wet chemistry treatment. The efficiency of the anticorrosion treatments were qualified using wide dispersive x-ray fluorescence (WDXRF) and chemical topography analyses by x-ray photoelectron spectroscopy (XPS) analyses. The WDXRF technique allows the overall chlorine level remaining after anticorrosion treatments to be determined throughout the volume of the metal stack. Using the combination of (1) geometrical shadowing of photoelectrons and (2) the differential charging of insulating portions of the features, the chemical topography¹⁰⁻¹³ of the metal features by XPS allows analysis of the chemical composition of the tops anti-reflective coating (ARC TiN), sidewalls (aluminum), and bottoms (SiO_2) of the metal features after anticorrosion treatments. By combining the results of WDXRF and XPS, the efficiency of various postetch anticorrosion treatments was assessed and the localization of the chlorine remaining on the metal surfaces could be accurately determined.

II. EXPERIMENT

Etching experiments were conducted using a LAM® Rainbow platform consisting of a load-lock chamber, etching chamber, DownStream Quartz (DSQ) installed in the exit load lock, and an integrated water rinse module. The etching chamber used was a LAM® transformer coupled plasma (TCP) equipped with an 8 in. coil and a half-shaped window. A process was developed for metal stack etching according to 0.25 μm design rules.

The anticorrosion treatment evaluated used the DSQ chamber. After etching, samples were transferred *in situ* into the DSQ chamber. The discharge was created in a quartz chamber by a TCP source. No bias was applied to the wafer during the DSQ treatment. The plasma treatment is a two-step etching process. In the first step, a high-pressure (1.2 Torr) H_2O plasma removes the chlorine from the resist and metal surfaces. The second step, performed at the same pressure, uses an $\text{O}_2/\text{H}_2\text{O}$ gas mixture (800 sccm/80 sccm) and acts as a resist removal step. During both steps, the TCP source power was 1000 W. After the DSQ treatment, the wafers were rinsed in the water rinse station.

After anticorrosion treatment using DSQ, a wet chemistry was used to remove chlorinated aluminum residues left on the metal stack sidewalls. The aqueous organic mixture used is commercialized by EKC Technology Inc. under the name of EKC 265®. The EKC 265 depolymerizer is composed of 2.2 aminoethoxyl ethanol (50%), hydroxylamine (35%), catechol (5%), and water. After etching and the DSQ anticorrosion treatment, the samples were immersed in an EKC bath heated to a temperature of 65 °C for 33 min, including rinsing and drying.

200 mm diam wafers were used. The following layers were deposited on Si wafers: 500 nm SiO_2 , 40 nm Ti, 60 nm TiN, 650 nm Al containing 0.5% Cu, 40 nm of TiN, and 1.1 μm of deep ultraviolet (DUV) photoresist (Shipley SNR

200®). The SNR resist was exposed using DUV light and developed with a resolution down to 0.25 μm .

The chlorine concentrations left on the metal sidewalls after anticorrosion treatments were measured using wide dispersive x-ray fluorescence and x-ray photoelectron spectroscopy. WDXRF combines the technique of x-ray fluorescence and diffraction of the emitted x-ray beam on a suitable single crystal.¹⁴ The crystal planes act as a monochromator grating for the fluorescent radiation emitted from the sample. By adjusting the incident angle between the beam and the crystal planes, the characteristic wavelength radiation of a specific element can be detected with excellent selectivity and sensitivity. Variations in stoichiometry between samples can be detected to within less than 0.8% of the element concentration in the volume of the sample. With a number of preconfigured detectors, specific different elements can be analyzed simultaneously, ranging from beryllium to uranium. Variations in stoichiometry measurements can be quickly obtained at various spots across the surface of the wafer allowing wafer mapping to be achieved. In this work, WDXRF was used to measure the chlorine level left in the metal lines after anticorrosion treatments. WDXRF allowed measurement of the chlorine level throughout the volume of the metal stacks after different treatments. However, the location of chlorine on metal lines could not be determined.

X-ray photoelectron spectroscopy experiments were conducted in an analysis chamber equipped with an x-ray photoelectron spectrometer (Fisons Surface Systems ESCALAB 220i), a single electrostatic lens flood gun (VG model LEG 41), and a mass spectrometer (Micromass 386). A monochromatized Al $K\alpha$ is provided for XPS analysis; further details on the analysis system can be found elsewhere.¹⁰ The source is provided with a source defined small spot toroidal monochromator. We used a 1000 μm spatial resolution mode resulting in an x-ray spot diameter of less than 1 mm.

The mask design used in this study allows characterization of the chemical topography of the tops, sidewalls, and bottoms of the features by XPS.¹⁰ Photoelectron signals from different portions of the features are separated using the methods of geometrical shadowing of photoelectrons and differential charging of the insulating and conducting surfaces. Lines and spaces can be arranged in parallel or perpendicularly to the axis of the electron analyzer. Consequently, the tops of the features and bottoms of the spaces (in the parallel mode) or tops and desired parts of the sidewalls, depending on the space between the lines, of the features (in the perpendicular mode) can be analyzed (see Fig. 1). The patterned areas analyzed in this study consist of regular arrays of trenches, a blanket substrate, and unpatterned mask material, the size of the areas being at least 1.5 mm^2 . This geometry allows the x-ray beam (1 mm diam) to be focused on one individual array, the resulting signal being the average of many identical lines and spaces. Variations in line and space widths allows analysis of the desired portion of the features. In this study, when the resist remained on top of the metal lines after metal etching and before the DSQ treatment, we analyzed the 1 μm line and space metal lines in the

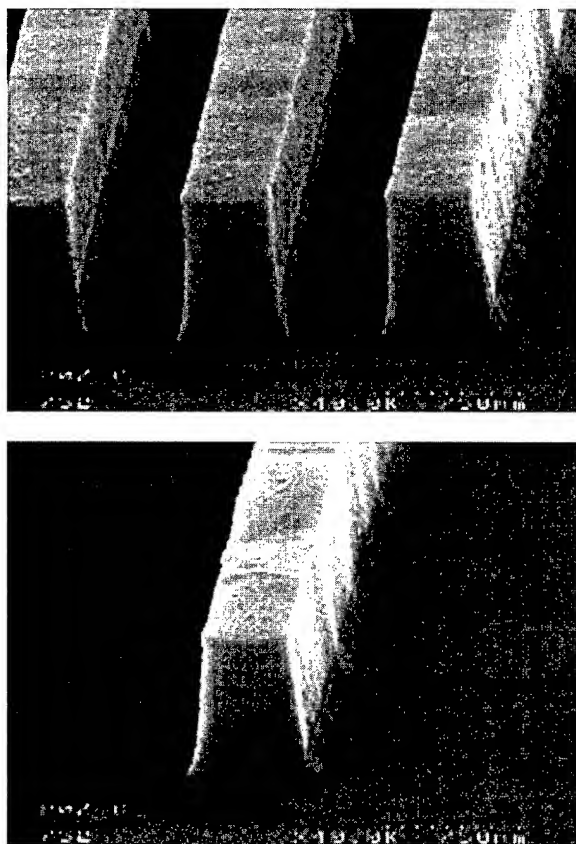


FIG. 1. Dense and isolated metal lines etched in the ICP source using a Cl_2/BCl_3 gas mixture.

parallel or perpendicular mode. When the resist was removed after metal etching and the DSQ anticorrosion treatment, we analyzed $0.5 \mu\text{m}$ line/space (L/S) structures (to use the shadowing effect and analyze only the aluminum sidewall in the perpendicular mode).

After etching and anticorrosion treatment, wafers were transferred *ex situ* into the analysis chamber. The wafer was placed in the focal plane of the electron energy analyzer and x-ray sources using a stage elevator. A charge coupled device video camera installed on top of the chamber was aligned with the focal point of the x-ray and electron energy analyzer and allowed selection of the desired area of the sample to be analyzed. Using this chemical topography analysis technique, chlorine concentrations were measured on the tops of the metal features (TiN), bottoms of the features (SiO_2), and sidewalls of the features (Aluminum). Contrary to WDXRF, XPS provided some information on the chemical composition of the near surface of the sample analyzed (usually, the thickness probed by XPS is around 10 nm).

The process developed in the LAM® TCP 9600 for metal etching according to $0.25 \mu\text{m}$ design rules was optimized using a statistical design of the experiments. The rf power injected into the source was 460 W and rf bias power 180 W. The gas mixture used was Cl_2/BCl_3 (65 sccm/20 sccm), the pressure in the etching chamber was fixed at 10 mTorr. The chuck temperature was controlled at 50°C throughout the

TABLE I. Chlorine level as measured by WDXRF in the volume of the metal stack after various treatments.

Metal stack processing	Chlorine concentration (arbitrary unit)
Reference sample	28
Metal etch (LAM TCP 9600)	1547
Metal etch + DSQ	50
Metal etch + DSQ + EKC	42

etching process. Figures 1(a) ($0.45 \mu\text{m}$ line, $0.6 \mu\text{m}$ space) and 1(b) ($0.45 \mu\text{m}$ isolated line) show that using these etching conditions followed by the DSQ treatment and wet depolymerizer EKC 265, anisotropic metal lines were obtained.

III. RESULTS

WDXRF experiments were performed to measure the chlorine concentrations on (a) a reference sample, (b) the metal stack after etching, (c) the metal stack after etching followed by the DSQ anticorrosion treatment, and (d) the metal stack after etching, the DSQ anticorrosion treatment, and the EKC 265 depolymerizer. The results are summarized in Table I. The reference was recorded before etching on a sample covered with the metal stack and photoresist patterns. WDXRF shows that after etching the metal stack, the chlorine level increases dramatically. The chlorine level measured on this sample averages the chlorine concentration present on the resist mask, on the sidewalls of the metal stack, and on the silicon dioxide at the bottom of the features. After etching and the DSQ anticorrosion treatment (which also acts as the resist removal step), the chlorine level decreases considerably (see Table I), but still remains at a higher level (approximately twice) than in the reference sample. This result shows that most of the chlorine measured in the sample after etching was trapped in the resist and that chlorine was still present after DSQ, either on the aluminum sidewalls of the features, or on SiO_2 at the bottom of the features (TiN on top of the features was not exposed to the Cl_2/BCl_3 plasma as it was protected by the photoresist mask during the etching process). After the entire process (metal stack etching + DSQ + EKC 265), the chlorine level decreases even more. As expected, the EKC wet depolymerizer

TABLE II. Element concentration measured by XPS in $1 \mu\text{m}$ L/S metal lines after metal etching. The lines are aligned perpendicularly to the electron energy analyzer.

Element after metal	Concentration (in %) etching
Al 2p	5.4
B 1s	...
C 1s	54.5
Cl 2p	29.0
F 1s	...
N 1s	...
O 1s	10.0
Si 2p	...
Ti 2p	1.1

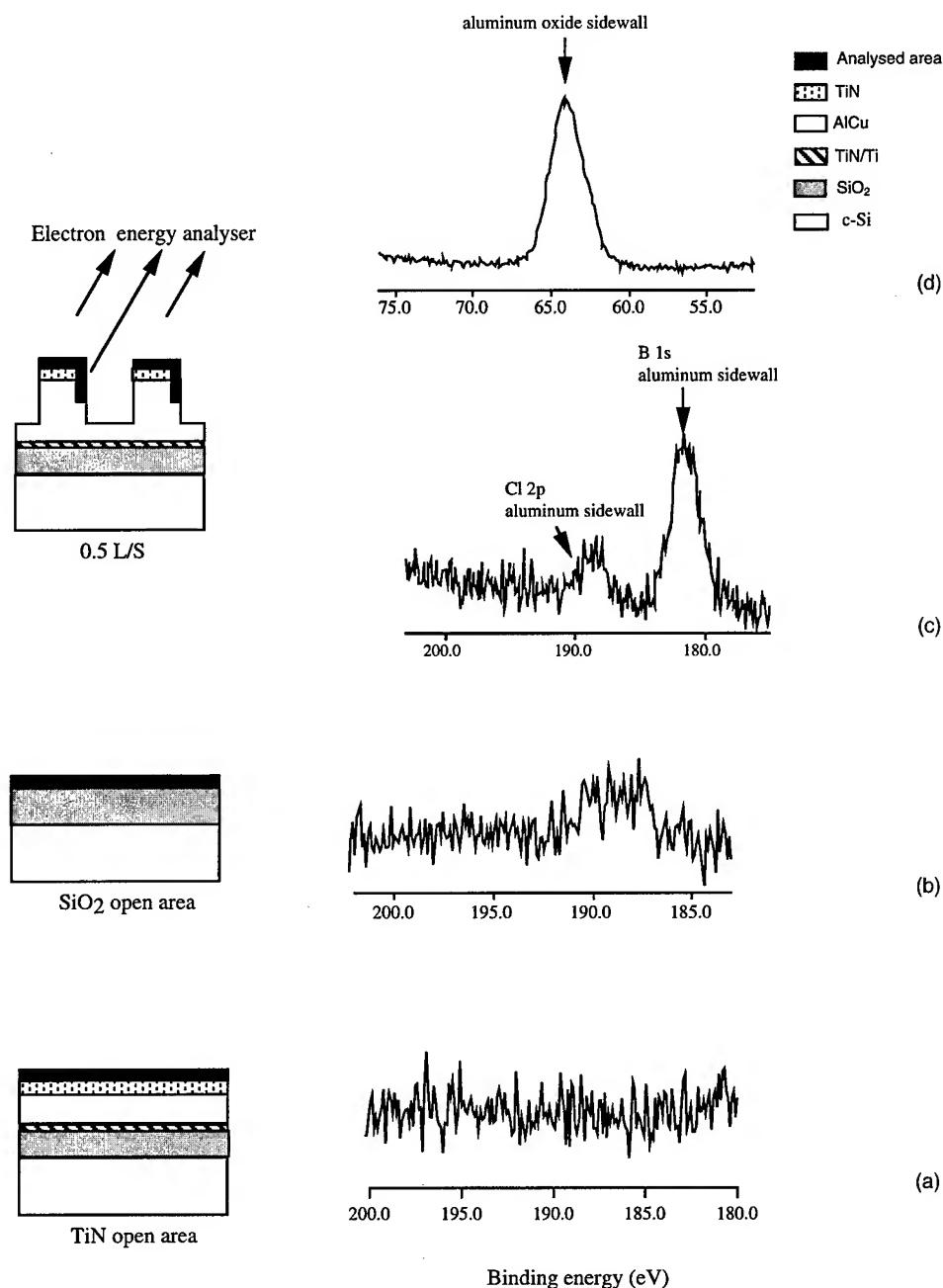


FIG. 2. Cl 2p XPS spectra recorded with the flood gun turned on after metal etching + DSQ on (a) blanket TiN, (b) blanket SiO₂, and (c) 0.5 μm lines and spaces aligned perpendicularly to the electron energy analyzer. (d) Al 2p XPS spectra recorded in 0.5 μm lines and spaces aligned perpendicularly to the electron energy analyzer.

removes a thin chlorine-containing aluminum layer. However, as the chlorine level is still higher than in the reference sample, the depolymerizer efficiency is not high enough to completely remove the chlorinated aluminum layer. These results show that chlorine originating from the etching process is still trapped in the aluminum lines and could possibly generate corrosion.

X-ray photoelectron spectroscopy analyses were conducted on the same samples. First, no chlorine was detected on the reference sample by XPS. After the etching process, XPS analyses recorded in 1 μm line and space metal lines

aligned in parallel with the electron energy analyzer (the analyzed areas are the top of the resist mask and SiO₂ at the bottom of the features) show that high concentrations of chlorine are still present on the resist mask (9%) and on SiO₂ at the bottom of the features (12.5%). XPS analyses were also performed in the same structures with the lines aligned perpendicularly to the electron energy analyzer (the analyzed areas are the tops and sidewalls of the resist mask and sidewalls of the metal lines). Concentrations of the elements (reported in Table II) were calculated by dividing the integrated peak intensities by the theoretical Scofield cross

TABLE III. Element concentration measured by XPS in 0.5 μm L/S metal lines after metal etching and DSQ treatment. The lines are aligned perpendicularly to the electron energy analyzer.

Element after metal	Concentration (in %) etching + DSQ
Al 2p	28.0
B 1s	5.2
C 1s	10.4
Cl 2p	0.4
F 1s	10.4
N 1s	1.4
O 1s	43.1
Si 2p	...
Ti 2p	1.1

sections (Al 2p:0.54, Si 2p:0.82, C 1s:1, Cl 2p:2.29, O 1s:2.93, F 1s:4.43, Ti 2p:7.91, N 1s:1.80, B 1s:0.48).¹⁵ When the charge neutralizer was turned on, no separation of the contributions coming from the resist mask and metal lines could be obtained. We attributed this inability to discriminate insulating and conducting surfaces to the thick carbon-based polymer deposited on all surfaces during the etching process, which electrically contact all surfaces. Consequently, the concentrations presented in Table II are an average of the element concentrations measured on the tops and sidewalls of the resist mask and aluminum sidewalls. Very high concentrations of carbon and chlorine were detected on the surfaces analyzed, whereas no elemental aluminum was detected on the sidewalls. These results show that the polymer, which builds up on the aluminum sidewalls and allowed the anisotropic etching profiles to be obtained, is a chlorine-rich carbon film. The absence of elemental aluminum originating from the sidewalls demonstrates that the carbon-based passivation polymer is thicker than the thickness probed by XPS. As the electron escape depth λ in a carbon-rich film can be estimated to be 35 Å,¹⁶ the sidewall thickness probed by XPS is $3\lambda \cos \theta$, i.e., 75 Å (θ is the angle between the normal to the sidewall and the electron energy analyzer). These results are in good agreement with XPS analyses performed *in situ* after etching wafers in another ICP commercial etcher with a Cl_2/BCl_3 gas mixture.⁷ XPS analyses also revealed that a chlorine-rich carbon film was formed on the sidewalls of the aluminum features during the etching process. The chemical sidewall composition indicated that chlorine-rich aluminum species were present in the passivation film. The chlorine-rich carbon polymer film enhances anisotropic etching by providing a thin protective layer against the spontaneous etching reaction of chlorine with aluminum.⁷ When the analyses were performed *in situ*, the chlorine concentrations measured on all surfaces were greater than the concentrations reported here. When the wafers are exposed to air after etching and before introduction into the XPS chamber, the surface of the sample is oxidized and the chlorine adsorbed on metal or resist surfaces may be replaced by oxygen atoms, therefore, leading to an underestimation of the true chlorine content in the film.

XPS analyses recorded on the sample after metal etching

and DSQ show that the chlorine concentration decreased dramatically on all surfaces of the metal stack. First, Cl 2p XPS spectra were recorded on blanket TiN [Fig. 2(a)] and SiO_2 [Fig. 2(b)] areas. In both cases, spectra were recorded with the flood gun on. As the shift in energy induced by the flood gun on TiN was 12 eV, Cl 2p should have been located at an energy of 187 eV. The shift in energy induced by the flood gun on SiO_2 was 5 eV. However, no Cl 2p was observed at a binding energy of 194 eV on the SiO_2 surface. Similarly, no chlorine was detected when the features were aligned in parallel with the electron energy analyzer (showing that there was no chlorine on TiN on the tops of the metal lines and no chlorine on SiO_2 at the bottoms of the features).

XPS spectra were also recorded with the flood gun turned on in 0.5 μm L/S structures aligned perpendicularly to the electron energy analyzer. As the tops and sidewalls of the features were conducting surfaces, it was impossible, even with the flood gun on, to separate the contributions originating from TiN on the top of the features and those from aluminum on the sidewalls. However, as the metal stack was formed on a 0.5 μm thick SiO_2 layer, the metal stack was electrically floating on SiO_2 and the flood gun induced chemical shifts. The presence of chlorine in the 0.5 μm L/S features demonstrated [see Fig. 2(c)], as there was no chlorine on the tops of the features, that all the chlorine-based species were located on the sidewalls of the aluminum stack (in this configuration, the analyzed areas are the tops and sides of the resist mask and aluminum sidewalls). An oxygen containing aluminum layer (called alumina oxide, hereafter) was detected on the aluminum sidewalls throughout the thickness probed by XPS [no elemental aluminum was detected on the sidewalls in Fig. 2(d)]. As a shift of 10 eV was induced by the flood gun on all the XPS peaks recorded in 0.5 μm L/S features, the Cl 2p peak was located at an energy of 189 eV, whereas a peak located at 182 eV was attributed to the B 1s present on the analyzed surfaces. The position of the Al 2p peak was 64 eV, showing that the Al 2p peak originated from an aluminum oxide layer formed on the aluminum sidewalls during the DSQ process.

Concentrations of the elements (reported in Table III) were calculated by dividing the integrated peak intensities by the theoretical Scofield cross sections.¹⁵ Consequently, the concentrations reported in Table III average the concentrations of the various elements detected on the analyzed area. A very high oxygen concentration (48%) was measured on the features as a logical consequence of the oxidation of the metal lines by the DSQ treatment. As aluminium is mainly linked to oxygen and as the electron escape depth λ of Al 2p photoelectrons in alumina is ~ 28 Å, the approximate depth probed by XPS on the sidewall was 60 Å ($3\lambda \cos \theta$), showing that the thickness of the aluminum oxide after the DSQ treatment was greater than 60 Å. Fluorine was also detected on the surfaces analyzed and was attributed to residual fluorine present in the DSQ chamber (due to previous treatments using CF_4/O_2 gas mixtures from other experiments). Carbon was detected on the surfaces (see Table III) and was attributed to contamination due to the transport of

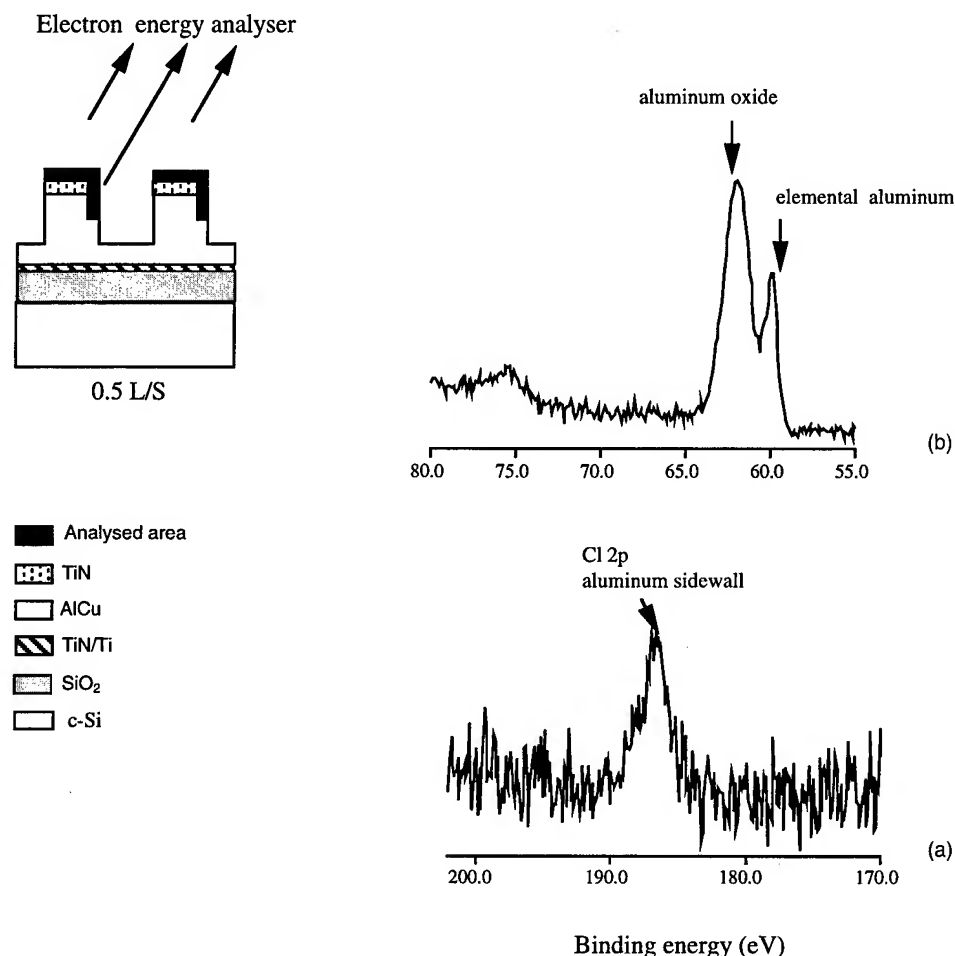


FIG. 3. XPS spectra recorded with the flood gun turned on in $0.5\ \mu\text{m}$ lines and spaces aligned perpendicularly to the electron energy analyzer after metal etching + DSQ + EKC (a) Cl $2p$ XPS spectrum, and (b) Al $2p$ XPS spectrum.

wafers in ambient air. As analyses in the parallel mode have shown that no chlorine was detected on TiN surfaces on the tops of the features, and as in $0.5\ \mu\text{m}$ L/S features the analyzed areas on the tops and sidewalls have the same size, the chlorine concentration on the aluminum sidewalls was twice as high as that measured (0.8%). XPS experiments demonstrate that the DSQ treatment efficiently removes chlorine from SiO_2 between the metal lines. They also show that during the DSQ treatment, an aluminum oxide film is formed on the sidewalls of the metal stack with chlorine still trapped inside.

XPS experiments were also conducted on samples after etching, DSQ, and depolymerization using the EKC 265. As above, no chlorine was detected on blanket TiN and SiO_2 surfaces and on TiN on the tops of the metal lines and on SiO_2 between the metal lines (when features were aligned in parallel with the electron energy analyzer). On the other hand, when XPS spectra were recorded with the flood gun on in $0.5\ \mu\text{m}$ L/S structures aligned perpendicularly to the electron energy analyzer, chlorine was also detected. Figure 3(a) shows that, due to charging, chlorine was shifted by 12 eV from its original position. As above, the detection of chlorine when metal lines are aligned perpendicularly to the electron

energy analyzer shows that chlorine-based species are located on the sidewalls of the aluminum stack. Elemental aluminum [see Fig. 3(b)] was detected on the sidewalls of the features (also shifted by 12 eV), showing that the thickness of alumina was thinner than that probed by XPS (less than 60 Å, based on the previous discussion). The calculated concentrations derived from the integrated peak intensities are listed in Table IV. A large amount of oxygen was present on all the surfaces analyzed (45% on average on the tops and sidewalls of the metal lines). On the other hand, no boron was detected and the fluorine concentration decreased from more than 10% to less than 2%. Carbon due to air contamination was still present on all surfaces. As no chlorine was detected on TiN on the tops of the features, the chlorine concentration on the aluminum surface originating from the sidewalls was, as above, twice as high as that measured, i.e., 1.8%.

IV. DISCUSSION

WDXRF and XPS experiments have shown that after metal etching, high concentrations of chlorine are present on all the metal stack surfaces exposed to high-density chlorine-based plasmas. Both techniques also show that the chlorine

TABLE IV. Element concentration measured by XPS in 0.5 μm L/S metal lines after metal etching, DSQ treatment, and wet depolymerizing chemistry. The lines are aligned perpendicularly to the electron energy analyzer.

Element after metal	Concentration (in %) etching + DSQ + EKC
Al 2p	23.5
B 1s	...
C 1s	17.0
Cl 2p	0.9
F 1s	1.5
N 1s	4.0
O 1s	46.0
Si 2p	...
Ti 2p	7.1

level is significantly decreased after the DSQ anticorrosion treatment. The DSQ treatment has two different roles. The first role is to act as a resist stripper and efficiently remove the photoresist in which a large amount of chlorine is trapped after metal etching. The second role is to remove the chlorine-rich carbon film, which builds up on the aluminum sidewalls during the etching process. After the DSQ treatment, XPS shows that chlorine is still present on the sidewalls of the metal lines. *In situ*⁷ and *ex situ* XPS analyses (this study) have also shown that chlorine-rich aluminum species are embedded in the sidewall passivation film after metal etching. Contrary to carbon-based species, which can easily be removed by the DSQ treatment, chlorine-rich aluminum species may be oxidized and trapped in the aluminum oxide film formed on the sidewalls during the Cl_2/BCl_3 etching process, possibly explaining why chlorine is still present after DSQ.

After metal etching + DSQ + EKC 265 depolymerizer, WDXRF analyses have shown that the EKC 265 has decreased the chlorine level measured in the volume of the stack (Table I) with respect to samples measured after metal etching + DSQ. These measurements confirm the role of the depolymerizer, which is expected to remove the chlorinated aluminum residues left on the metal stack. However, as the chlorine level is still higher than in the reference sample, the depolymerizer did not remove all the chlorine containing aluminum film. XPS analyses confirm that EKC has removed a thin aluminum sidewall film as, again, elemental aluminum was detected by XPS (showing that the aluminum oxide film thickness has decreased on the sidewall). On the other hand, XPS analyses have shown that after DSQ + EKC 265, the chlorine concentration on the aluminum sidewalls was greater than after DSQ, which seems to be in contradiction with the WDXRF results showing that throughout the volume of the metal stack, the chlorine level has decreased.

A plausible hypothesis explaining the difference between the WDXRF and XPS measurements is that (1) DSQ does not lower the chlorine concentration throughout the thickness of the chlorine-containing film left on the sidewall of the stack and (2) EKC thins or may even totally remove the layer, which has been depleted in chlorine by the DSQ, but not by all the chlorine-containing metal layer. The composi-

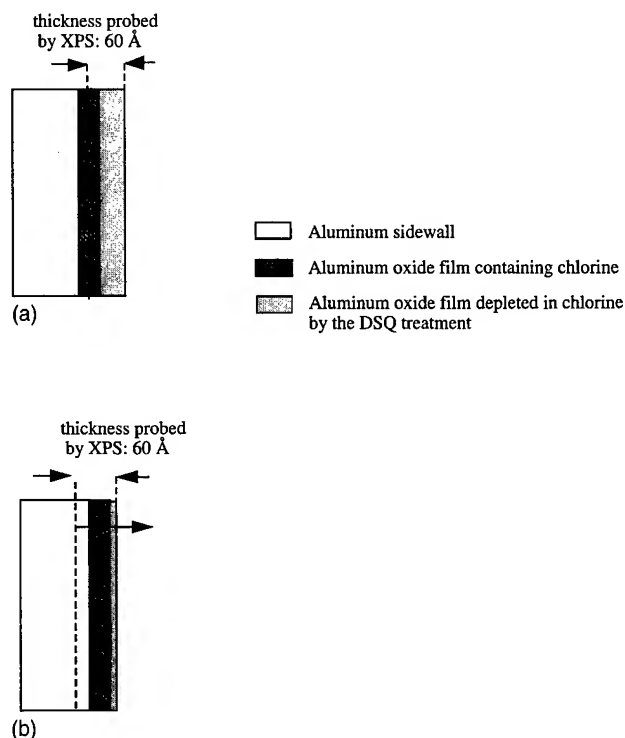


FIG. 4. Schematic view of the aluminum sidewall surface composition (a) after metal etching + DSQ, and (b) after metal etching + DSQ + wet depolymerizer.

tion of the aluminum sidewall corresponding to the mechanism proposed above is represented in Fig. 4. Figure 4(a) is a schematic view of the aluminum sidewall after metal etching + DSQ. Figure 4(b) is a schematic view of the aluminum sidewall after metal etching + DSQ + EKC. After DSQ, a first layer on the surface of the aluminum sidewall is depleted in chlorine and a second layer contains a higher chlorine content [see Fig. 3(a)]. As XPS is more sensitive to the top surface, the XPS Cl 2p signal mostly originates from the chlorine-depleted layer. On the other hand, after DSQ + EKC, EKC has partially removed the chlorine-depleted layer and XPS probes a layer containing a higher chlorine concentration. The chlorine concentration as measured by the intensity of the XPS Cl 2p peak is, therefore, higher than after DSQ, whereas the chlorine level in the volume of the stack has, indeed, decreased (as measured by WDXRF). In Fig. 4(b), we illustrate the case in which EKC has partially removed the chlorine-depleted layer by DSQ. However, EKC possibly removes all the chlorine-depleted layer. In this case, XPS will still measure a higher chlorine concentration on the aluminum sidewalls than after DSQ. Furthermore, the schematic view represented in Figs. 4(a) and 4(b) represents the case in which the chlorine concentrations in both layers are uniform, whereas a gradient in chlorine would probably be more representative of the real case, which cannot be readily measured. However, the conclusions based on the discussion above would again be unchanged.

Combining the WDXRF technique, which is sensitive to the chlorine level variation in the volume of the metal stack and XPS analyses, which are more sensitive to the surface

modifications, is of great help in understanding the mechanisms involved in the anticorrosion treatments. Based on the discussion above, the chlorine level on the sidewalls of the aluminum lines could be decreased even more by reducing the passivation sidewall polymer thickness formed during metal stack etching or by reducing the chlorine content of this film. This could be achieved by developing a process using plasma operating conditions, producing a thinner sidewall passivation film while maintaining the anisotropy of the etching. The DSQ treatment may also have a significant impact on the final chlorine level in the metal line. Indeed, although DSQ removes all the chlorine-rich carbon species from the aluminum surfaces, it may also favor the diffusion of chlorine deeper into the aluminum (during DSQ the substrate temperature is 250 °C). If, after DSQ treatment, chlorine species have been trapped in the aluminum metal lines due to diffusion, wet depolymerizing chemistries become ineffective at removing them. Therefore, reducing the DSQ temperature could also be an interesting way of decreasing the chlorine level in the metal stack. A compromise would have to be found between the chlorine diffusion in the metal line and the efficiency of DSQ at removing chlorine-rich carbon species from the aluminum sidewalls. Finally, improving the efficiency of wet depolymerizing chemistries could also improve the overall efficiency of anticorrosion treatments.

V. CONCLUSION

Chemical topography analyses by XPS combined with WDXRF have been used to evaluate the efficiency of widely used metal stack anticorrosion treatments consisting of high-pressure water-based plasmas followed by depolymerizing wet chemistry. Prior to anticorrosion treatments, wafers were etched in a low-pressure ICP plasma using Cl_2/BCl_3 gas mixtures. The most important result is that, after metal etching followed by the complete anticorrosion treatment, residual chlorine is still present in the metal stack, showing that corrosion could occur in the subsequent technological steps of the process. Chemical topography analyses by XPS have shown that, after anticorrosion treatment, residual chlo-

rine species are located on the aluminum sidewalls of the features. Combining the results obtained by WDXRF and XPS has led to a better understanding of the action of both the DSQ anticorrosion treatment and the wet depolymerizer. The DSQ lowers the chlorine concentration at the near surface of the aluminum sidewall but not throughout the chlorinated aluminum layer. The wet depolymerizer partially or totally removes the layer, which has been depleted in chlorine by DSQ, but is not efficient enough to remove the all chlorine-containing metal layer.

ACKNOWLEDGMENTS

This work has been carried out within the framework of the GRESSI consortium between CEA-LETI and France Telecom-CNET. It is part of the Adequat+ Esprit program.

- ¹Y. Teraoka, H. Aoki, I. Nishiyama, E. Ikawa, and T. Kikkawa, *J. Vac. Sci. Technol. A* **13**, 2935 (1995).
- ²C. Gabriel and R. Wallach, *Proceedings of the Symposia on Reliability of Semiconductor Devices/Interconnections and Dielectric Breakdown for Microelectronic Applications*, 1992, p. 152.
- ³T. Ishida, N. Fujiwara, M. Yoneda, K. Nakamoto, and K. Horie, *Jpn. J. Appl. Phys.* **31**, 2045 (1992).
- ⁴S. Madokoro, S. Hagiwara, and S. Ushio, *Extended Abstracts of the 16th Conference on Solid State Devices and Materials*, Kobe, Japan (1984), pp. 51–54.
- ⁵G. Cameron and A. Chambers, *Semicond. Int.* **12**, 142 (1989).
- ⁶C. B. Labelle, H. L. Maynard, and J. T. C. Lee, *J. Vac. Sci. Technol. B* **14**, 2574 (1996).
- ⁷P. Czaprynski, O. Joubert, L. Vallier, M. Heitzmann, and P. Berruyer, *J. Vac. Sci. Technol. B* (submitted).
- ⁸F. R. Clayton and S. A. Beeson, *Solid State Technol.* **36**, 93 (1993).
- ⁹S. Bradley, C. H. Chen, and G. Kovall, *Proceedings of the Eighth International IEEE VLSI Multilevel Conference*, Santa Clara, CA, 1991, p. 298.
- ¹⁰F. H. Bell, O. Joubert, and L. Vallier, *J. Vac. Sci. Technol. B* **14**, 1796 (1996).
- ¹¹M. C. Peignon, F. Clenet, and G. Turban, *J. Electrochem. Soc.* **143**, 1347 (1996).
- ¹²G. S. Oehrlein, K. Chan, and M. Jaso, *J. Appl. Phys.* **64**, 2399 (1988).
- ¹³K. V. Guinn and V. M. Donnelly, *J. Appl. Phys.* **75**, 2227 (1994).
- ¹⁴A. Reader and K. Nieuwenhuizen, *Solid State Technol.* **16**, 516 (1996).
- ¹⁵J. H. Scofield, *J. Electron Spectrosc.* **8**, 129 (1976).
- ¹⁶S. Tanuma, C. J. Powell, and D. R. Penn, *Surf. Interface Anal.* **11**, 577 (1988).

Isolation of a lattice-mismatched AlInAs/GaInAs layer on InP using ion implantation for high energy mobility transistor realization

Hervé Fourré,^{a)} Jean Claude Pesant, Olivier Schuler, and Alain Cappy

Institut d'Electronique et de Microelectronique du Nord U.M.R. C.N.R.S. 9929, Département Hyperfréquences et Semiconducteurs, Cité Scientifique, 59652 Villeneuve d'Ascq, Cedex, France

(Received 16 January 1997; accepted 30 May 1997)

In this communication, we present Ar⁺ implant isolation results of InGaAs/InAlAs/InP layers for high electron mobility transistor fabrication. Ar⁺ implantations were performed at 100 keV at room temperature with doses ranging from 5×10^{12} to 10^{15} at./cm². A sheet resistance of about 30 M Ω was measured after 10 hours annealing at 300 °C. Using deep Auger analysis, we show that this high resistivity is due to the intermixing of InGaAs and InAlAs layers. © 1997 American Vacuum Society. [S0734-211X(97)15004-1]

I. INTRODUCTION

InGaAs/InAlAs/InP material based transistors are being used increasingly for millimeter wave integrated circuits and high speed applications. For large scale production of high electron mobility transistors (HEMTs) on an InP substrate, ion implantation is preferred to mesa etching for device isolation, for several reasons.

- (1) Implant isolation maintains a planar structure, which is desirable in device fabrication and wafer processing.
- (2) Implant isolation avoids breaking the gate at the mesa sidewalls.
- (3) Implant isolation suppresses the gate leakage problem related to contact of the gate with the highly doped channel layer through the sidewall.

Implant isolation of InGaAs/InAlAs/InP HEMT layers is, however, difficult. The energy band gap of the InGaAs composition, which is lattice matched (LM) on InP, is only 0.74 eV and the intrinsic resistivity of this material is small. We can expect difficulties in achieving good isolation in this material by ion implantation. On the other hand, a high-resistivity region can be obtained by implantation in InAlAs lattice matched on InP. Sheet resistances of about $10^7 \Omega$ have been achieved this way in LM-InAlAs after implantation of either He⁺ (Ref. 1) or O⁺ (Ref. 2) ions, while the highest sheet resistance obtained with LM-InGaAs after implantation is only about $10^5 \Omega$. Unfortunately, this value is too low for device applications. In a previous article, we presented implant isolation of an InP LM-HEMT layer using H⁺, B⁺, F⁺, and Ar⁺ ions.³ Proton implantation leads to very poor isolation of only 3 k Ω /square and this result correlates well with previously reported H⁺ implantation in InGaAs (Refs. 2 and 4) that shows a sheet resistance of only 500 Ω /square after implantation. B⁺ or F⁺ implantation of-

fers much better isolation and our best results were obtained by implanting either 10^{14} B⁺/cm² at 20 keV or 5×10^{15} F⁺/cm² at 35 keV. In both cases, post-implant annealing was performed at 300 °C for 10 minutes and the sheet resistances obtained were, respectively, about 1 M Ω /square and 2 M Ω /square. These results are close to the value reported by Fink *et al.*⁵ using O⁺ implant isolation for LM-HEMTs.

According to the sheet resistance variation as a function of annealing time and temperature, we have deduced that B⁺ and F⁺ implant isolation is due to damage-induced defects.⁶ Ar⁺ implantation also provides good results with regard to these different ions. We chose Ar⁺ because of its heavier ion mass (40) compared to B⁺ (10) and F⁺ (18). We can expect that this ion will, therefore, generate more defects in the layer. In this communication we present the results of Ar⁺ implantation investigation and analysis. Particular problems associated with the fabrication of LM-HEMT are addressed and discussed. Results obtained with Ar⁺-isolated transistors are also presented.

II. EXPERIMENTAL PROCEDURE

Different implantations were performed using an Eaton Ga 3204 which allows room temperature implantation at 200 keV with an ion current of 3 mA. The structure of the epilayers used in this work is shown in Figure 1. In order to avoid channeling effects, the samples were implanted with an incident angle of 7° normal to the sample surface. Investigations of Ar⁺ implantation were performed for doses ranging from 5×10^{12} to 1×10^{15} at./cm². The expected range (Rp) for these implantations is 100 nm for a 100 keV implant energy. This value is found in the InAlAs buffer layer (Figure 1). The first sample (Figure 1(a)) was used for sheet resistance measurements by transmission line methods (TLM) patterns, which were fabricated prior to the implantations. The samples's initial sheet resistance was 230 Ω /square. The non-implanted regions were protected by a

^{a)}Electronic mail: fourre@helena.iemn.univ-lille.fr

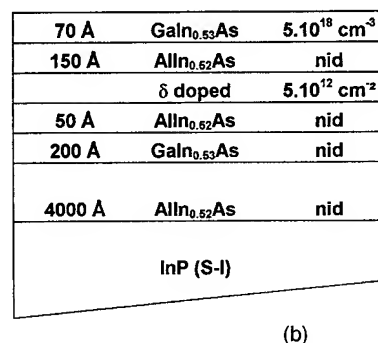
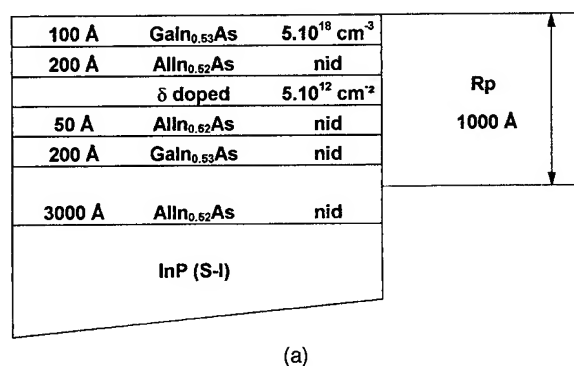
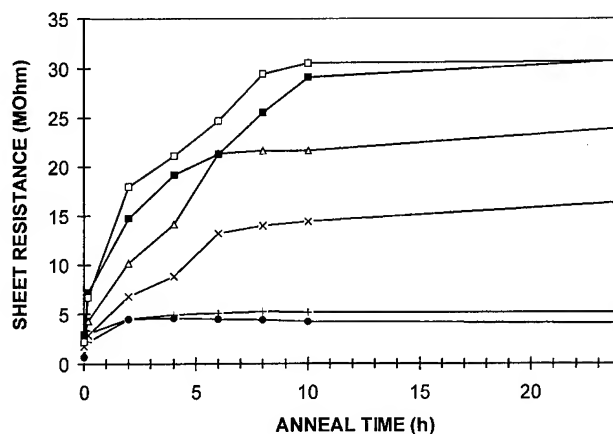


FIG. 1. (a) and (b) Structure of the implanted epilayers.

thin layer of Si_xN_y deposited by plasma-enhanced chemical vapour deposition (PECVD) at 300 °C, and a thick photoresist mask. It should be noted that a lift-off profile for the resist mask was used because it greatly reduces the polymer traces after wafer cleaning. After removing the photoresist mask, post-implant annealing was performed in a tubular furnace under N₂ gas for times ranging from 10 minutes to 24 hours at 300 °C. Samples for Auger analysis were implanted and annealed under the conditions that gave the best isolation without any protection. The layer used for Auger analysis is shown in Figure 1(b).

The implant and annealing parameters were carefully chosen to be compatible with our HEMT fabrication process. One major difficulty in the implant isolation process is to find some good protection for the non-implanted regions that can afterwards be removed while keeping the wafer clean. During our experiments, we found that a 1.4 μm thick resist mask offered good protection. But ion bombardment leads to resist degradation and leads to polymer formation. As a consequence, the epilayer is left with traces of polymers after removing the resist. The resist mask degradation is related to both current density and total ion dose. We have found that a lift-off profile for the resist mask reduces the number of polymer traces after acetone cleaning. For high current densities, the sample temperature increases and more polymer is created. Hence, we chose a current density of about 0.25 μA/cm². Moreover, since the total amount of polymer created is proportional to the implanted ion dose, the ion dose was less than 10¹⁵ Ar⁺/cm² throughout this work. In addition, a 20 nm Si_xN_y layer was deposited by PECVD at 300 °C prior to the resist mask deposition. This thin Si_xN_y

FIG. 2. Sheet resistance of a 100 keV Ar⁺ implanted HEMT layer vs annealing time at 300 °C for different doses (+) 5×10¹², (●) 1×10¹³, (×) 5×10¹³, (Δ) 1×10¹⁴, (■) 5×10¹⁴, and (□) 1×10¹⁵ cm⁻².

layer allowed us to perform O₂ reactive plasma etching (RIE) to remove traces of the polymer without degradation of the non-implanted areas which correspond to the active layer of the HEMT. The chosen post implant annealing temperature was about 300 °C in order to avoid degradation of the TiAu alignment marks used in the electronic lithography process. No experiments were performed at higher annealing temperatures.

III. RESULTS AND DISCUSSION

Figure 2 shows the sheet resistance of Ar⁺ implanted samples measured for different doses and different annealing times at 300 °C (from 0 to 24 hours). The highest sheet resistance measured was about 30 MΩ/square, for implanted ion doses greater than 5×10¹⁴ cm⁻² and for annealing times greater than 10 hours at 300 °C. In order to investigate the isolation mechanisms, sputtering Auger electron analysis was carried out for a non-implanted sample and for a 1×10¹⁵ cm⁻² implanted sample annealed at 300 °C for 10 hours. Figure 3 shows the deep Auger profiles for these two samples. For the non-implanted sample, the different layers are clearly defined. In the second Auger profile, we observe the diffusion of Ga and Al species and the formation of InGaAlAs layers. In In_{1-x-y}Ga_xAl_yAs layers, sheet resistances greater than 10 MΩ/square have already been reported with He⁺ implantation.⁷ The sheet resistances obtained in this work are close to this value.

Therefore, the isolation mechanisms for Ar⁺ implants can be explained as follows. Due to their heavy mass, Ar⁺ ions create many defects in the layer during the implantation process. During annealing, these defects lead to intermixing, by diffusion. This diffusion process is well correlated with the annealing time dependence of the sheet resistance shown in Figure 2. Moreover, this behavior is in agreement with the annealing time and annealing temperature dependence reported for InGaAs/InP intermixing.⁸ It should be noted that

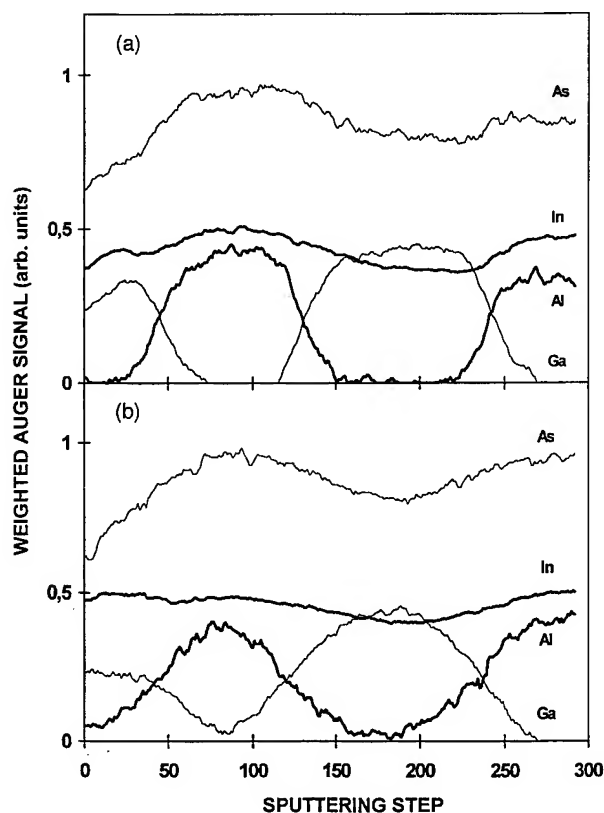


FIG. 3. Depth Auger profile of HEMT layer: (a) as grown, (b) implanted (1×10^{15} Ar^+/cm^2 , 100 keV) and annealed (300 °C 10 hours).

the annealing temperature is too low to remove ion damage since the lowest temperature at which ion damage removal was reported is 500 °C.^{6,9}

It should be emphasized that our implant isolation process characterized by room-temperature implantation and low-temperature annealing requires thin InGaAs layers and comparatively thicker InAlAs layers, which are the Al source for the InGaAlAs layer. Other implantation processes, in which intermixing phenomena have been observed, have been already reported. In particular, it is well known that ion implantation performed at a temperature higher than a critical T_c promotes intermixing.^{6,8-10} Furthermore, in this implantation technique no post-implant annealing is necessary. For III-V materials the values of T_c is generally between 200 °C and 300 °C. This implantation process also has the advantage of being compatible with HEMT fabrication. However, it has been suggested that the number of defects generated by implantation at temperatures above T_c become depleted after implantation.⁸ Actually, our implantation equipment does not allow us to perform high temperature implantation, and again we do not know if the number of defects would be sufficient and efficient enough to provide good isolation in the inter-

mixed InGaAlAs layers.

IV. DEVICE APPLICATION

To demonstrate that ion implantation is a promising HEMT processing technique, we fabricated implant-isolated LM-HEMTs on InP. The epilayer used to fabricate these implanted HEMTs is given in Figure 1(b), except for the channel layer which was a 15 nm thick $\text{In}_{0.65}\text{Ga}_{0.35}\text{As}$ pseudomorphic layer. Prior to implantation, the sheet resistance was 170 Ω/square . After implantation, it was measured by TLM patterns to be about 49 $\text{M}\Omega/\text{square}$. This is the highest resistivity obtained in this work and we assume that it is due to the presence of a pseudomorphic InGaAs layer. We know that Ar^+ implantation can create defects that favor strain relaxation by formation of misfit dislocations. The implant-isolated transistors have electrical performance similar to that of the mesa-isolated transistors fabricated on the same layer. This indicates that the proposed implantation process did not degrade the active part of the devices.

V. CONCLUSION

Implant isolation of InP LM-HEMT layers using Ar^+ ions was investigated. Good isolation was obtained with sheet resistances greater than 30 $\text{M}\Omega/\text{square}$ using moderately high doses (10^{15} cm^{-2}) and long-term annealing at relatively low temperature (300 °C, 10 hours). To our knowledge, this is the best isolation result for LM-HEMT on an InP layer obtained by implantation. This implantation process is compatible with LM-HEMT fabrication. The feasibility of HEMT using this process was demonstrated.

ACKNOWLEDGMENT

The authors would like to thank Jean Pierre Nys for performing the Auger spectra.

- ¹W. Lee and C. G. Fonstad, *Appl. Phys. Lett.* **50**, 1278 (1987).
- ²*Indium Phosphide and Related Materials: Processing, Technology, and Devices*, edited by A. Katz (Artech House, Boston, MA, 1991), pp. 222-239.
- ³H. Fourré, O. Schuler, J. C. Pesant, A. Leroy, and A. Cappy, *Indium Phosphide and Related Materials Conference Proceeding*, 1996, p. 331.
- ⁴B. Tell, K. F. Brown-Goebeler, T. J. Bridges, and E. G. Burkhardt, *J. Appl. Phys.* **60**, 665 (1986).
- ⁵T. Fink, B. Raynor, M. Haupt, K. Köhler, J. Braunstein, N. Grün, and J. Hornung, *J. Vac. Sci. Technol. B* **12**, 3332 (1994).
- ⁶S. J. Pearton, C. R. Abernathy, M. B. Panish, R. A. Hamm, and L. M. Lunardi, *J. Appl. Phys.* **66**, 656 (1989).
- ⁷B. Tell, T. Y. Chang, K. F. Brown-Goebeler, J. M. Kuo, and N. J. Sauer, *J. Appl. Phys.* **64**, 290 (1988).
- ⁸W. Xia, S. A. Pappert, B. Zhu, A. R. Clawson, P. K. L. Yu, and S. S. Lau, *J. Appl. Phys.* **71**, 2602 (1992).
- ⁹W. Xia, S. C. Lin, S. A. Pappert, C. A. Hewett, M. Fernandes, T. T. Vu, P. K. L. Yu, and S. S. Lau, *Appl. Phys. Lett.* **55**, 2020 (1989).
- ¹⁰S. A. Pappert, W. Xia, B. Zhu, A. R. Clawson, Z. F. Guan, P. K. L. Yu, and S. S. Lau, *J. Appl. Phys.* **72**, 1306 (1992).

Quantitative measurement of two-dimensional dopant profile by cross-sectional scanning capacitance microscopy

J. S. McMurray, J. Kim, and C. C. Williams^{a)}

Department of Physics, University of Utah, Salt Lake City, Utah 84112

(Received 7 January 1997; accepted 4 April 1997)

Quantitative two-dimensional dopant profiling of a gatelike structure is achieved by scanning capacitance microscopy (SCM). A processed silicon wafer is sectioned then polished such that a cross section is made through a gatelike structure. This structure consists of two n+ doped regions separated by a lighter doped n region. The two-dimensional SCM data are converted to dopant density through a physical model of the SCM-silicon interaction. Improvements to the physical model and SCM data to dopant profile algorithm are discussed. Advances in sample preparation are described. Adjustment of the conversion model parameters allows for fitting of the SCM dopant profile to a vertical secondary-ion-mass-spectroscopy profile. The accuracy of this fit is better than 20%. This fit also gives strong evidence that the full two-dimensional profile is truly quantitative.

© 1997 American Vacuum Society. [S0734-211X(97)00404-6]

I. INTRODUCTION

Quantitative two-dimensional (2D) dopant profiling is identified as one of the significant needs of the semiconductor industry.¹ Having 2D dopant information provides for calibration of technology computer aided design (TCAD) simulators and will shorten process development time.² Several methods are being investigated for 2D dopant profiling. These include: selective etching, 2D secondary-ion-mass-spectrometry (SIMS), Nano-spreading resistance profiling (SRP), and scanning capacitance microscopy (SCM) techniques. These technologies are under development, and have been recently reviewed by Diebold *et al.*³ The dopant profiling performed by SCM uses a conversion model to convert SCM data into 2D dopant profiles. Quantitative vertical dopant profiles from SCM measurements have previously been reported.⁴ This paper represents the first 2D dopant profile of a 2D structure, quantitatively converted directly from SCM data. The 2D profiles are a result of improved sample preparation, better physical modeling of the SCM-silicon interaction, and an improved SCM data to dopant profile conversion algorithm.

II. TECHNIQUE AND SAMPLE PREPARATION

The SCM technique uses an atomic force microscope (AFM) to scan a conducting tip over a sample. The tip is electrically connected to a sensitive capacitance sensor. Local changes in the sample's capacitance are measured as the tip is scanned. In this way, both topographic and capacitance data are acquired simultaneously. For the measurements discussed, a NanoScope IIIa AFM, with a Dimension 3000 head manufactured by Digital Instruments, is used. The capacitance measurements are performed with an RCA capacitance sensor.^{5,6} In this configuration, the input capacitance to the RCA sensor is about 1 pF. Most of this is stray capacitance. This sensor uses a 915 MHz probing voltage for capacitance detection. To separate the small tip capacitance from the

much larger stray capacitance, it is necessary to look for changes in capacitance. This is accomplished on semiconducting samples by applying a time varying bias voltage to the sample. The bias voltage typically has a frequency between 5 and 15 kHz. This bias voltage modulates the depletion capacitance in the semiconductor. This is a standard SCM configuration which has been successfully used with both commercial and home built AFMs.^{7,8}

For this system to work, it is necessary to have an insulating layer between the tip and semiconductor. It is found that the ~2 nm native oxide that grows on silicon is too thin.⁹ The voltages required to operate the SCM cause conduction through this oxide. A thicker oxide (~3 nm or greater) is required for the SCM measurement. To have maximum capacitance change, the bias voltage applied to the sample must be able to drive the silicon under the tip into both deep depletion and accumulation. Measured capacitance versus voltage (*C-V*) curves show that this voltage can be upwards of 4 V peak.¹⁰ While large bias voltages are good for large signals, they are not good for high resolution. The SCM spatial resolution is determined by the volume of the silicon that is depleted under the tip. With a large bias amplitude and lightly doped sample, this depleted region is larger than the tip diameter. Resolution degradation can be overcome by varying the amplitude of the applied bias in such a way that the change in capacitance measured by the sensor is small and is held constant.¹¹ This leads to large bias voltages in heavily doped regions where the depletion is kept small by the large number of carriers and a small bias in lightly doped regions. Using this configuration, the SCM resolution is limited by the diameter of the tip.

The amplitude of the bias voltage is measured and related to the dopant density through a conversion algorithm. The algorithm is based on a quasi-3D model of the tip sample capacitor.⁴ Using this method, high resolution vertical dopant profiles on silicon wafers have been achieved.^{4,12} These profiles are found to be in good agreement with vertical SIMS measurements.

The previous results report only vertical profiles obtained

^{a)}Electronic mail: clayton@physics.utah.edu

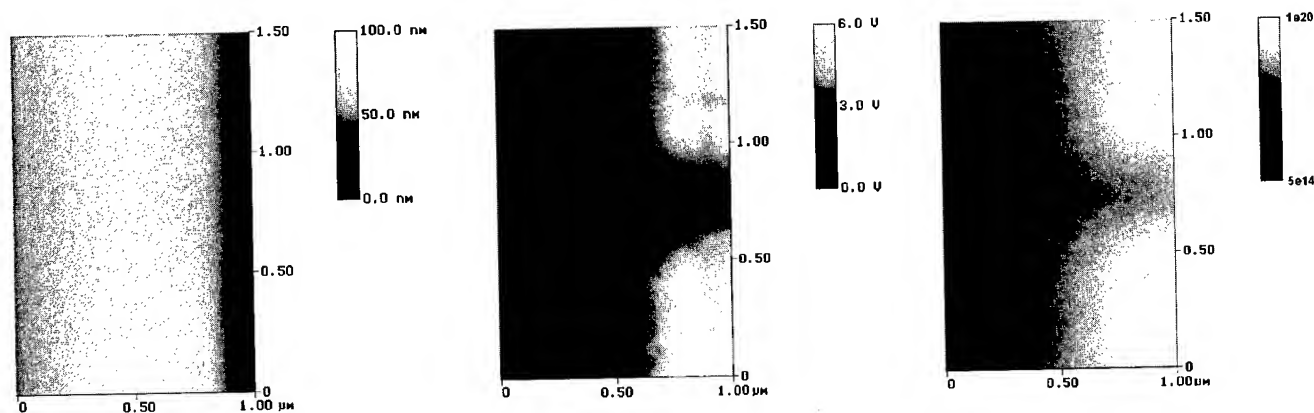


FIG. 1. Cross-sectional SCM data showing a gatelike structure. Image (a) shows the topographic data, (b) the applied depletion bias, and (c) the converted dopant density.

on cross-sectioned surfaces by SCM. Advances in sample preparation and the conversion algorithm have been made in order to obtain full 2D quantitative profiles. A large SCM signal is needed to allow for fast data acquisition. One way to achieve this is to use thinner oxides on the silicon surface. Previously, cross-sectional samples were cleaved and then a sputtered SiO_2 layer was deposited on the surface. This method presents two problems for continued use: the thickness of the very thin oxides (5 nm) is hard to control in the sputtering system, and cleaving is too imprecise a process to allow for examination of specific structures on a silicon wafer. To provide for examination of structures, the sample can be cleaved near the area of interest and then polished to the structure using standard techniques. It is found that for the SCM data to be of adequate quality, topographic height variations caused by polishing must be less than 1 nm over a $1 \mu\text{m}^2$ region.

In the final step of the polishing process, a colloidal silica suspension solution is used.¹³ Samples prepared in this way require no additional deposition of oxide on the surface for insulation. The colloidal silica solution contains a hydroxide that reacts with the silicon during polishing to produce a thin insulating layer. The size of the C - V curves generated with the SCM indicates that the layer on the surface after polishing is thin (about 3 nm) and of sufficient insulating quality to be useful for SCM. However, it is observed that sometimes the C - V curves are offset by more than 5 V. We believe this is due to trapped charge at the surface. Since this insulating layer is quite thin and uniform over the surface it is not expected to affect the dopant distribution.

III. MEASUREMENT AND CONVERSION TO DOPANT DENSITY

The 2D profile reported here is obtained on an n-type sample that is implanted with 80 keV phosphorus ions with a dose of 10^{15} cm^{-2} . The substrate doping level is approximately $10^{15} \text{ atoms cm}^{-3}$ and the peak concentration is $8 \times 10^{19} \text{ atoms cm}^{-3}$. The sample is prepared, as described, by cleaving and polishing to the desired implanted area. The area of interest is a series of gatelike structures; heavily im-

planted regions separated by a lighter doped region underneath the gates. These structures are created by ion implantation, with $0.56\text{-}\mu\text{m}$ -wide lines of oxide used to mask the gate area. Figure 1 shows a $1.5 \mu\text{m}$ cross-sectional scan of this gatelike structure. The first image (a) is the topography. The sample edge is seen on the right side of this image. The edge is not abrupt due to slight rounding that has occurred during polishing. The second image (b) is the amplitude of the applied depletion bias to the sample. Note that the bias is greatest in the heavily implanted regions. The third image (c) is the inverted 2D dopant profile. The effect of the topographic edge is seen in image (b) as a drop in the depletion bias voltage. All data in images (b) and (c) to the right of the topographic edge are meaningless. Figure 2 shows cuts through the 2D profile in Fig. 1. These cuts show the dopant density variation due to the oxide mask. The data collected has 512 points per line and 128 lines per image. The 128 lines are interpolated to 512 lines. The data are processed with a 16 pixel Gaussian filter in both the x and y

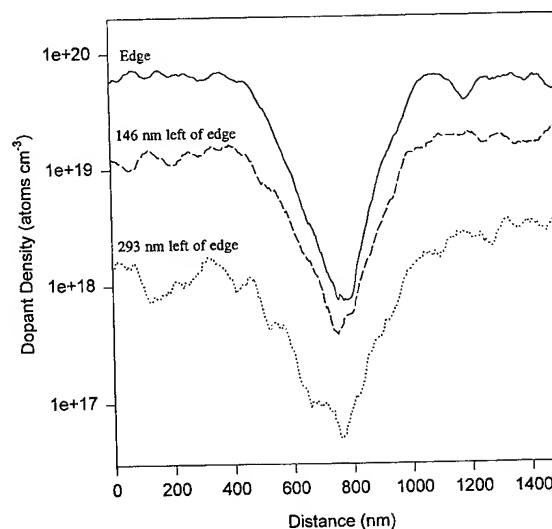


FIG. 2. Line cuts from image (c) in Figure 1, taken parallel to the topographic edge through the gatelike structure.

directions before the conversion. The Gaussian width used in the filter was 25 nm which is less than the tip's radius. The Gaussian filter is used to average out noise in the measurement. Since the filter size is kept small relative to both the tip and the dopant gradient no information is lost in this process.

This 2D dopant profile is achieved with several improvements to the physical model and the conversion algorithm.¹⁴ In the earlier model, the tip was represented by a sphere, embedded in a half plane of dielectric. The model now uses an approximate solution for a tip surrounded with air, resting on a thin sheet of oxide which is covering the silicon. Changes to the conversion algorithm have resulted in it being faster and more accurate. The root finding procedures have been improved. The iteration to find the dopant density for each data point is replaced by a simulation of the required sample bias for a range of dopant densities. The values are stored in a look up table. The conversion is then accomplished by simply matching a measured bias for each data point with the corresponding value in the look up table and performing an interpolation to find the desired dopant density. These changes allow the improved conversion algorithm to convert a 2D array of 512×512 SCM data points to a 2D profile in 10–20 min on a 90 MHz Pentium computer, more than 300 times faster than the original algorithm. The variance in the conversion time is mainly due to the modeling of the sensor probing voltage and is almost independent of the number of points in the input data.

The conversion algorithm requires several parameters. These are: tip radius, peak dopant density and corresponding ac bias, oxide thickness, oxide dielectric constant, and size of the SCM sensor probing voltage. In this measurement a heavily doped silicon tip with a radius of 37 nm (determined by imaging with niobium film¹⁵) was used. The peak dopant value used in this conversion is 8×10^{19} atoms cm^{-3} , and is determined by SIMS. Changing any of the other model parameters (oxide thickness, dielectric constant, or the amplitude of the sensor probing voltage) has a similar effect; they all affect the lowest dopant value generated with the conversion process. Currently none of these parameters is exactly known, but reasonable values can be chosen for two of the three, and the third varied as a free parameter within reasonable limits. The two values fixed in this conversion are the oxide thickness and the oxide dielectric constant. The thickness is chosen to be 3 nm, slightly larger than a native silicon oxide. The value of the dielectric constant used is 3.0. This value was chosen to be similar to a chemical vapor deposition (CVD) oxide dielectric constant.

The value of the free parameter, the sensor probing voltage, is adjusted so that the dopant profile taken in the vertical direction in an implanted region most closely fits the SIMS profile. The final value for the sensor probing voltage is 1.15 V peak, a reasonable value. Figure 3 shows two SCM dopant profiles compared with a vertical SIMS profile, and the corresponding topographic information from the SCM data. The two SCM profiles are taken horizontally from the 2D profile shown in Fig. 1. The finely dotted line is extracted from the 2D profile 30 nm from the top of image (a) in Fig. 1 and the

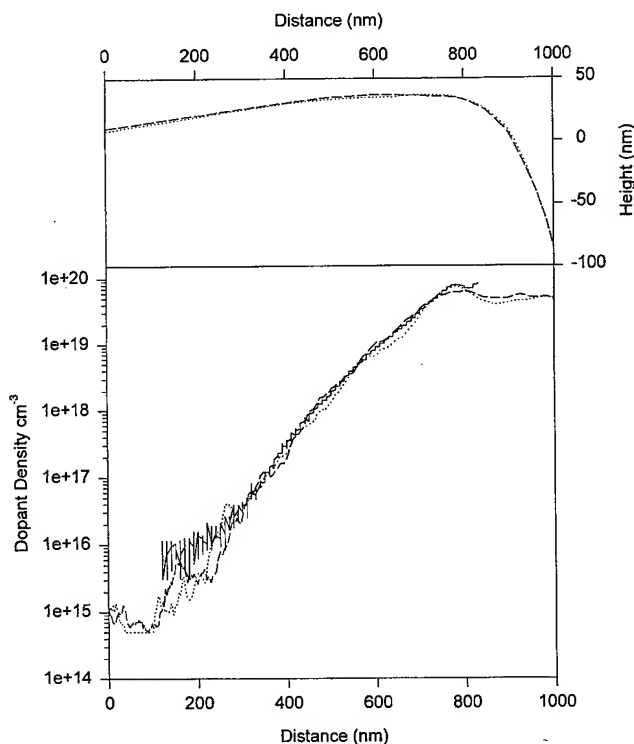


FIG. 3. (a) SCM topographic data and (b) a comparison of SCM vertical profiles to a SIMS vertical profile.

dashed line 30 nm from the bottom. Since the sample's edge is not well defined, the SCM and SIMS curves are shifted laterally for the best fit. The shift from the apparent edge is not more than ± 15 nm.

The fit of the dopant profile to the SIMS profile is quantified by taking a point by point comparison, calculating the absolute percent difference for each point, and averaging the result. The dopant range considered is 10^{20} – 10^{17} atoms cm^{-3} . The average percent difference for the dashed line is 17%, while for the dotted line it is 12%. The same model parameters are used for all the points in converting the 2D SCM data and there is no preferential direction in the data acquisition. However, the conversion model assumes the dopant density is uniform underneath the tip. This can lead to errors in the accuracy of the converted profile if the change in dopant density is large across the diameter of the tip and if the gradient varies significantly across the sample. In this case, the dopant change is not large (one order of magnitude for 200 nm) and the tip is small (37 nm radius). In addition the dopant change is approximately the same whether taken vertically or laterally through the gatelike region. Therefore, any effect the dopant change causes should be approximately the same throughout the sample, and is compensated for by fitting to the vertical SIMS profile. This indicates that the accuracy in the full 2D image should be comparable to that of the vertical profile.

IV. CONCLUSION

Quantitative 2D dopant profiling is performed by SCM. This is achieved with improved sample preparation, im-

proved physical modeling of the SCM-silicon system, and an improved conversion of SCM data to dopant profiles. The accuracy of the SCM dopant profile is compared with a SIMS profile of the same sample and is in good agreement in the vertical direction. This provides strong evidence that the full 2D profile is quantitative. Future work will concentrate on devices with junctions and on achieving higher resolution results.

ACKNOWLEDGMENTS

The authors would like to thank Digital Instruments, SEMATECH, and SRC for their help and funding of this project, and J. Slinkman at IBM for providing the sample.

¹The National Technology Roadmap for Semiconductors is published by the Semiconductor Industry Association, 4300 Stevens Creek Boulevard, Suite 271, San Jose, CA 95129.

²M. E. Law, *Simulations of Semiconductor Devices and Processes*, edited by S. Selberherr, H. Stippel, and E. Strasser (Springer, Berlin, 1992), Vol. 5, p. 1.

³A. C. Diebold, M. R. Kump, J. J. Kopanski, and D. G. Seiler, *J. Vac. Sci. Technol. B* **14**, 196 (1996).

⁴Y. Huang, C. C. Williams, and H. Smith, *J. Vac. Sci. Technol. B* **14**, 443 (1996).

⁵R. C. Palmer, E. J. Denlinger, and H. Kawamoto, *RCA Rev.* **43**, 194 (1982).

⁶J. K. Clemens, *RCA Rev.* **39**, 33 (1978).

⁷J. J. Kopanski, J. F. Marchiando, and J. R. Lowney, *J. Vac. Sci. Technol. B* **14**, 242 (1996).

⁸C. C. Williams, Y. Huang, and K. S. Mak, Proceedings of the American Physical Society March Meeting, Indianapolis, Indiana, 16–20 March, 1992, postdeadline Session.

⁹G. Neubauer, A. Erickson, C. C. Williams, J. J. Kopanski, M. Rodgers, and D. Adderton, *J. Vac. Sci. Technol. B* **14**, 426 (1996).

¹⁰Y. Huang, C. C. Williams, and H. Smith, *J. Vac. Sci. Technol. B* **12**, 369 (1994).

¹¹Y. Huang, C. C. Williams, and J. Slinkman, *Appl. Phys. Lett.* **66**, 344 (1995).

¹²Y. Huang, C. C. Williams, and M. A. Wendman, *J. Vac. Sci. Technol. B* **14**, 1168 (1996).

¹³A. Erickson, Digital Instruments, Santa Barbara CA (private communications).

¹⁴Y. Huang, Ph.D. thesis, Physics Dept., University of Utah, 1995.

¹⁵K. L. Westra, A. W. Mitchell, and D. J. Thompson, *J. Appl. Phys.* **74**, 3608 (1993).

**Papers from the
24th Conference on the Physics
and Chemistry of Semiconductor Interfaces**

**12–15 January 1997
Sheraton Imperial Hotel
and Convention Center
Research Triangle Park, NC**

**Sponsored by
American Vacuum Society
Army Research Office
Office of Naval Research**

**Special Editor for the Proceedings:
David E. Aspnes**



**Published for the American Vacuum Society by
the American Institute of Physics, New York, 1997**

ORGANIZING COMMITTEE

PROGRAM COMMITTEE

L. Cooper (Office of Naval Research)
J. D. Dow (Arizona State University)
R. M. Feenstra (Carnegie-Mellon University)
D. K. Ferry (Arizona State University)
K. L. Kavanagh (University of California-San Diego)
G. Lucovsky (North Carolina State University)
T. C. McGill (California Institute of Technology)
C. G. Van de Walle (Xerox Palo Alto Research Center)
D. J. Wolford (Iowa State University)

CHAIR

D. E. Aspnes (North Carolina State University)

LOCAL ARRANGEMENTS CHAIR

J. E. Rowe (Army Research Office)

Conference Secretary
Carole Dow

Local Secretary
Rebecca Savage

PREFACE

This volume contains papers presented at the 24th Annual Conference on the Physics and Chemistry of Semiconductor Interfaces (PCSI-24). The Conference was held at the Sheraton Imperial Hotel and Convention Center, Research Triangle Park, NC, 12–15 January 1997. The Center provided a relaxing environment for the meeting, with the schedule, poster sessions, and social events arranged to encourage interaction.

The Conference has continued to emphasize fundamental theoretical and experimental understanding of interface problems relevant to emerging areas of technology. The size of the meeting, as measured by the number of attendees and abstracts submitted, was about the same as in 1996. A total of 75 papers were presented in 8 primary categories of topics, including invited talks intermixed with oral and poster presentations. In keeping with the Workshop tradition, every invited and oral presentation was accompanied by a poster. This allowed abundant opportunities for in-depth discussions during the informal break as well as the scheduled poster sessions.

Topics presented can be divided approximately into four major and four minor categories. With the usual caveat that some papers could be placed in several categories, the breakdown is as follows. The major categories were: silicon and related materials (16 papers); wide-bandgap materials, mainly nitrides and organics (15 papers); heteroepitaxy (14 papers), and nanostructures and self-assembly (12 papers). Minor categories included metal–semiconductor interfaces (6 papers); surface and interface spectroscopy (6 papers); interface analysis (4 papers); and passivation (2 papers). Thus the topics of nanoscale, self-assembly, and wide-bandgap materials continued to gain attention. The traditional Tuesday evening rump session focused on proximal probes with the provocative title “Proximal probes: what good are they?” Application of lateral resolution to improve our understanding of fundamental processes is expected to be a primary theme of PCSI-25 which is scheduled to meet in Utah. Decreasing activity was evident in metal–semiconductor interfaces and in passivation, although the Sunday evening talk by Professor Karl Hess on device applications of deuterium passivation received considerable attention.

The conference succeeded through the efforts of many individuals. We would like to recognize in particular the many anonymous reviewers who performed their work carefully, thoroughly, and efficiently, ensuring that the high standards of the Journal of Vacuum Science and Technology B were maintained; Jack Dow for conference management; Carole Dow for registration management; Rebecca Savage for local management; and Becky York for providing a smooth interface between the PCSI Proceedings Editor and the Journal of Vacuum Science and Technology B. The commitment of Gary McGuire for making this collection a part of JVST B is also appreciated, along with that of Galen Fisher, who provided an effective link with the American Vacuum Society. Financial support from Larry Cooper of the Office of Naval Research and Jack Rowe from the Army Research Office is gratefully acknowledged. The conference was held under the sponsorship of the American Vacuum Society through the Electronic Materials and Processing Division, the Office of Naval Research, and the Army Research Office.

Looking ahead, PCSI-25 will be held 18–22 January 1998 in Salt Lake City and will be chaired by R. M. Feenstra. (feenstra@andrew.cmu.edu; <http://stm1.phys.cmu.edu/pcsi/>).

D. E. Aspnes
Conference Chair and Proceedings Editor

J. E. Rowe
Local Arrangements Chair

Atomic wires and their electronic properties

Toshishige Yamada^{a)}

ERATO Quantum Fluctuation Project, Edward L. Ginzton Laboratory, Stanford University, Stanford, California 94305-4085

(Received 12 January 1997; accepted 23 March 1997)

Atomic wire electronics are considered, where the band structure and the resultant Fermi energy are designed by manipulating the lattice constant. Using the tight-binding theory with universal parameters, it is shown that Si wires and arrays are metallic, Mg wires are insulating, and Mg arrays have metallic and insulating phases for infinitely large, isolated cases. Structures are of finite size, and electrodes are necessary for the applications. The finite size brings about discrete electron energy levels, and electrodes will charge or discharge the structure, reflecting the work function difference, so that even the basic electronic properties may be altered. The electrodes will cause further complications such as the energy level broadening, the Coulomb interaction through an effective capacitance, or the mode-selection resistance. When the contact satisfies certain conditions, a metallic wire is predicted to show two distinct I - V patterns for small voltages. Depending on whether the highest occupied level is either partially or fully filled, the current starts to flow rapidly or does not flow until the voltage overcomes the next level, respectively. © 1997 American Vacuum Society. [S0734-211X(97)08704-0]

I. INTRODUCTION

Due to the recent progress in scanning tunnel microscope (STM) technology,¹ it is now becoming possible to manipulate atoms, or place them at desired positions. It is also routinely possible to prepare an atomically flat surface possessing two-dimensional periodicity. One such example is a reconstructed semiconductor surface. Using the STM atom manipulation technology, foreign atoms can be placed at the minima of the substrate periodic potential to form a one-dimensional atomic wire or a two-dimensional atomic array. After successful electronic isolation between the atomic structure and the substrate, electrons are confined within the structure and a low-dimensional electron system is obtained. The most attractive feature of this system is that we can manipulate the lattice constant and design the band structure and the resultant Fermi energy.²

The choice of a substrate providing both two-dimensional potential and electronic insulation is not trivial and we are studying the possibility of using a Si(100)2×1 reconstructed surface where the dangling bonds are saturated with hydrogen atoms.^{3,4} Foreign atoms are placed at the minima of the periodic potential created by substrate surface atoms via van der Waals-like interparticle interaction, and an atomic structure such as a wire is created. The mechanical stability of such a wire may be an experimental concern: there might be a tendency for the wire atoms to deform the wire shape spontaneously, e.g., clustering or dimerization. This issue can be addressed with a phenomenological interparticle potential.^{5,6} The potential includes the effects of kinetic, Coulomb, and exchange energies, and is featured in general by a hard core barrier for short distances, followed by a potential well that gives the natural separation of the particles, with an exponentially decaying tail for longer dis-

tances. In the present case, the separation among constituent wire atoms is chosen to be well beyond the natural distance and close to or inside the tail region, while the constituent atoms and the substrate atoms keep the natural distances, so that the wire will be mechanically stable.

The electronic properties of various atomic structures with Si or Mg atoms are studied⁴ with the tight-binding calculation with universal parameters,⁷ assuming that the structures are successfully isolated from the substrate and are infinitely large. It is shown that one-dimensional Si wires and two-dimensional Si arrays are metallic regardless of the lattice constant. One-dimensional wires of Si diatomic molecules are also metallic. All of these results reflect the fact that the highest occupied level is $3p$ and is partially filled in an isolated Si atom, and the Fermi energy is located in the $3p$ -originating band. It seems quite difficult to obtain a semiconductor or an insulator with Si atoms having low-dimensional atomic structures. We thus need to use group II atoms such as Mg or Be atoms for a semiconductor or an insulator, since the s level is fully filled and the p levels are completely empty in these atoms, and their periodic structures will have a fully filled lower band and an empty upper band. According to the calculation, Mg wires are insulating or semiconducting, and two-dimensional Mg arrays have metallic and insulating phases depending on the lattice constant.

Metallic structures with partially filled bands will be conductive and insulating structures with completely filled bands will not be conductive, as long as the response to a small electric field is concerned. One may further expect that the metallic wires will exhibit an integer multiple of a quantum unit of conductance in the linear response regime since there are at most several one-dimensional modes (discrete momenta in the transverse directions and continuous momenta in the longitudinal direction) involved in the wire transport. In fact, the present situation is similar to that of mesoscopy,

^{a)}Electronic mail: yamada@loki.stanford.edu

where the conductance quantization in small structures possessing at most a few modes has been verified theoretically and experimentally;⁸ this is independent of the details of the dispersion (how the discrete energy levels form in the finite wire limit).

This naive view is subject to various modification.⁹ First, electrodes have to be placed at the ends of the atomic structures for electronic measurements and device applications. Because of the possible work function difference between the wire and the electrodes, there would be a space-charge layer formation with an atomic scale, and the entire or part of the wire will be charged.^{9,10} Once the wire is charged, the Coulomb interaction might be important in the transport, via the effective capacitance at the contact between the wire and the electrodes. Second, the number of atoms used in a wire is limited to an order of 10^1 as long as the structure is created, e.g., by placing foreign atoms one by one with STM atom manipulation technology. Due to the finite wire length, the discreteness of the energy levels, or a zero-dimensional effect (discrete momenta in the transverse and longitudinal directions) in the wire will be relevant.

The physical processes at the contact decide the relative importance of the Coulomb interaction and the discreteness of the energy levels.⁹ When the particle exchange between the wire and the electrodes is extremely rare, an effective capacitance at the contact will be extremely small and the Coulomb interaction will be relevant. The longitudinal quantization effects are quite important, but the Coulomb interaction significantly modifies the discrete energy levels through the effective capacitance and even the spin degeneracy will be lifted. As the particle exchange rate increases, the Coulomb interaction is less and less relevant and the discrete energy levels are more and more broadened. When the particle exchange is extremely frequent, the electrons spread over the entire system and the effective capacitance is extremely large. The discrete nature of electron energy level is no longer present, and continuous bands appear.

With intermediate particle exchange, where the capacitance is large and the discrete nature of the electron levels is still relevant, we can expect two distinguished transport patterns reflecting how the electrons are accommodated in these discrete levels. When the highest occupied level of the wire is fully filled, the current does not flow without a voltage greater than the separation of the highest occupied and lowest unoccupied levels. Or, there is an offset voltage for the current onset. When the highest occupied level is partially filled, the current can flow even with a voltage smaller than the scale of quantized energy separation, but the differential conductance will not be infinite, but limited on the order of the quantum unit of conductance, due to the mode-selection resistance at the contact as in the mesoscopic systems.⁸ Each time the Fermi energies in the electrodes cross a new wire level, there will be a sudden increase in current, resulting in stairlike I - V characteristics.

In Sec. II, the band properties of the Si and Mg atomic structures are shown, using the tight-binding model with universal parameters. In Sec. III, physical processes at the con-

tact between the wire and the electrodes are discussed. In Sec. IV, I - V characteristics for metallic wires with two different fillings for the highest occupied level are shown, reflecting the momentum quantization along the wire. Summary and discussions are given in Sec. V.

II. ATOMIC WIRE ELECTRONICS-ELECTRONIC PROPERTIES OF Si AND Mg STRUCTURES

As long as the system has periodicity, we can find the following general behavior of the electronic states when changing the lattice constant without changing the lattice symmetry.² This general behavior can apply for all the lattice symmetry and dimensions.¹¹ When the lattice constant is large, the electrons are localized in each atom. The energy levels are in principle discrete and equivalent to those of isolated atoms, e.g., with an s level and three p levels in addition to the stable core levels that have rare gas configurations. As the lattice constant is reduced, the discrete energy levels have a finite width to form a band, because of an overlap of neighboring electron wave functions. At this stage, there are one lower state and three upper states, and each state can accommodate two electrons with spin up and down. If the lattice constant is further reduced, the band widths become so wide that one of the p -originating upper bands and the s -originating lower band touch. This is called band crossing. With further reduction in the lattice constant, there would be three possible cases, depending on the lattice symmetry and the dimension. (1) The touching upper band sticks to the lower band and they behave together. The resulting band structure is such that there are two lower states and two upper states, separated by a forbidden gap. Due to the change in accommodation number in the lower and upper bands, this will provide a mechanism for a metal-insulator transition. The usual diamond-structure Si crystal is semiconducting since the natural lattice constant happens to correspond to the one after band crossing. (2) The touching upper and lower bands repel each other to reopen a band gap after band crossing. There is no change in the number of states in the upper and lower bands, although the band gap is modulated with the lattice constant, resulting in no metal-insulator transition. This is the case for one-dimensional atomic wires. (3) The touching upper and lower bands overlap after band crossing, and the overlapped region is allowed, with each energy region corresponding to momenta in different directions. Thus, a large unified allowed energy region appears. This will also provide a mechanism for a metal-insulator transition. This behavior can be seen for two-dimensional arrays.

By using the tight-binding theory with universal parameters,⁷ it is shown that Si wires are metallic and Mg wires are insulating regardless of the lattice constant.⁴ Figures 1 and 2 show the band structure of Si and Mg wires for three different lattice constants, respectively. In principle, we can see the general behavior of the electronic states corresponding to (2). The only band gap reduces as the lattice constant is decreased, and vanishes at 2.7 Å for the Si wire and at 4.2 Å for the Mg wire. These lattice constants corre-

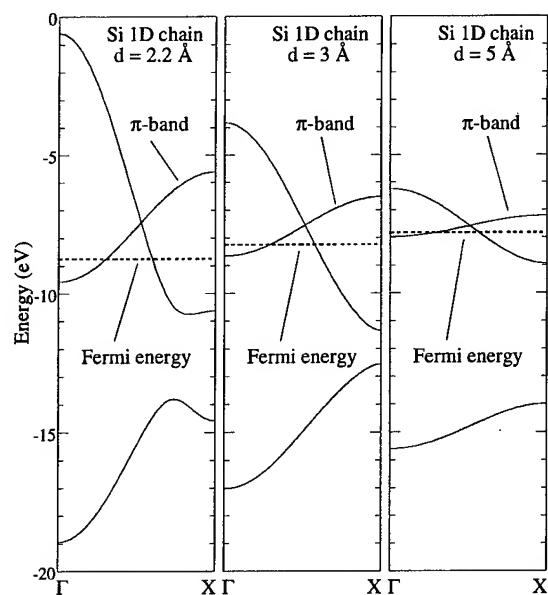


FIG. 1. Band structure of Si wire for selected lattice constants.

spond to band crossing. After this band crossing, a band gap reopens in both wires and we again have one state in the lower band and three states in the upper bands. The number of electrons that can be accommodated in the lower and upper bands remains the same. The major difference between Si and Mg is the number of electrons, or the Fermi energy position. In an isolated Si atom, the highest occupied level is $3p$ and is one-third filled. Since the number of states in the lower and upper bands remains the same regardless of the lattice constant, the Fermi energy always lies within the π band originating from two $3p$ levels normal to the wire and

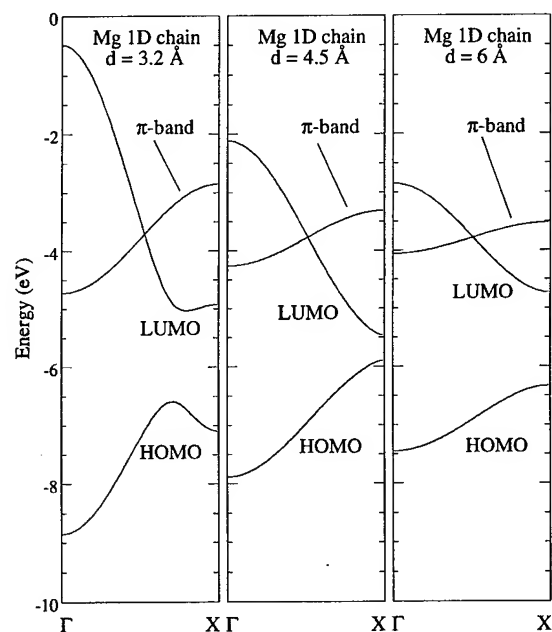


FIG. 2. Band structure of Mg wire for selected lattice constants.

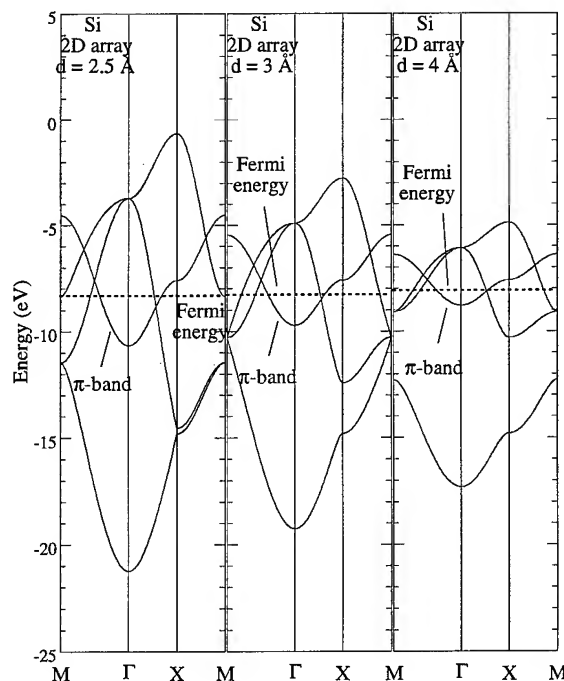


FIG. 3. Band structure of Si array for selected lattice constants.

this makes the Si wire metallic. In an isolated Mg atom, the highest occupied level is $3s$ and is fully filled while the $3p$ levels are empty. The lower band is always fully filled and the upper band is always completely empty, regardless of the lattice constant. This is why Mg wires are insulating or semi-conducting.

Figure 3 shows that two-dimensional Si square arrays are metallic regardless of the lattice constant d , for the same mechanism as Si wires, where the Fermi energy is again within the π band. The only band gap disappears at $d = 3.4$ Å after band crossing, corresponding to the general behavior of (3) above, but the Fermi energy is always well above the band gap and its disappearance has no effect on the electronic properties. If we can create a Si diatomic molecule so that the highest occupied molecular level is fully filled, then a Si insulator would be possible by simply placing these Si diatomic molecules along a line. It turns out that this idea does not really work since the required interatomic distance is unreasonably small, less than the natural Si atom diameter, but it would work for C atoms.⁴ Thus, it seems quite difficult to obtain an Si insulator having one- or two-dimensional lattice structures. Two-dimensional Mg square arrays have metallic and insulating phases depending on the lattice constant d , reflecting the disappearance of the band gap as shown in Fig. 4. For $d > 5.2$ Å, the lower band is fully filled and the upper bands are empty, resulting in an insulator. At $d = 5.2$ Å, the band crossing occurs and the band gap disappears. For $d < 5.2$ Å, the array is metallic since the upper and lower bands overlap and a large unified allowed energy region appears.

There is a pioneering experiment towards atomic wire electronics by Strosio, Feenstra, and Fein.¹² They have ob-

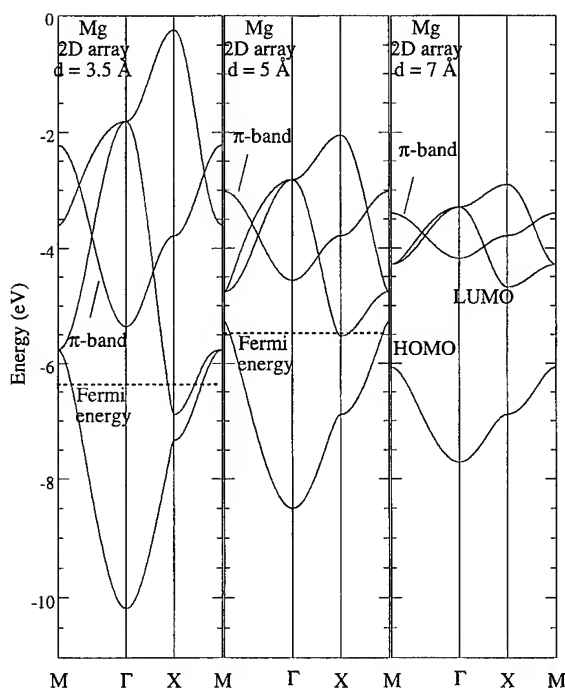


FIG. 4. Band structure of Mg array for selected lattice constants.

served a cleaved Si(111)2×1 surface with STM. The surface is characterized by a one-dimensional zig-zag wire of Si atoms with two basis atoms in a unit cell of 6.65 Å × 3.84 Å. Practically, we may think that one-dimensional wires consisting of dangling bonds with the lattice constant of 3.84 Å are placed in parallel, and the separation of wires are so large (6.65 Å) that interwire coupling is negligible. If the neighboring dangling bonds successfully couple in the wire direction and the electron wave spreads coherently over the entire wire, upper and lower bands are expected to form, which are separated by the only band gap. The electron filling is such that the lower band is fully filled and the upper band is empty,^{12,13} so that the Fermi energy lies around the middle of the band gap and the wire is insulating. The positive substrate voltage (shifting the entire substrate bands downward with respect to the tip Fermi energy) is expected to image the upper empty band and the negative voltage the lower filled band. At each Si position, the electron wave function has a maximum for the lower band states and has a minimum for the upper band states (the strongest contrast at the Brillouin zone edge), thus creating π -phase shift difference between the two images of opposite bias polarity.

They have experimentally observed this π -phase difference in STM images for positive and negative substrate voltages. This experiment demonstrates a coherent wavefunction overlap between neighboring atom sites in the wire direction even with a separation of ~ 4 Å, much wider than the natural Si atom separation of 2.35 Å in the usual diamond-structure crystal. The overlap is still strong enough for electrons to form a band structure in spite of this large lattice constant. This is encouraging for the present situation. The lattice constant of an atomic structure has to be chosen

well beyond the atom diameter to achieve a reasonable mechanical stability in order to avoid the possible deformation such as clustering or dimerization, but this may cause unwanted decoherence. Their experiment shows that the coherency is quite robust, and the decoherence would not be the case.

If we proceed to use foreign atoms to create an atomic structure, an appropriate substrate providing both periodic potential and electronic insulation has to be found. One possibility¹⁴ is to use a Si reconstructed surface where the surface dangling bonds are saturated with hydrogen atoms to form a monohydrogen layer.^{3,4} The dangling bonds are saturated so that the surface is expected to provide a two-dimensional van der Waals potential plane, and wire atoms are placed at the bottom of this potential without creating chemical bonds with the substrate atoms. The constituent wire atoms keep distances of well beyond their atom diameter, while the wire atoms and substrate atoms have distances corresponding to the natural separation.

There might be concern about the possible deformation such as clustering or dimerization after aligning wire atoms along a line, and this problem could be addressed with the help of a phenomenological interparticle potential.^{5,6} The potential includes the effects of kinetic, Coulomb, and exchange energies and is featured by a hard core barrier for small distances, and a shallow potential well outside the barrier that gives the natural distances, followed by an exponentially decaying, slowly varying potential tail for large distances. The detail of the functional form depends on the kind of atoms. In the case of Si-Si interparticle-interaction, it is shown⁶ that the potential minimum is at $\sim 2-2.5$ Å (the natural diameter of the Si atom) with the depth of ~ 2 eV, followed by an exponentially decaying tail for the distances larger than $\sim 3-3.5$ Å with a negligible depth. The clustering of a Si wire will be relevant when the lattice constant is small, roughly corresponding to $2-2.5$ Å. For the lattice constant to be larger than or equal to $3-3.5$ Å, the deformation will be unlikely. The wire atoms are held by substrate atoms with the natural distance, while the wire atoms themselves keep distances well beyond the natural distance, so that the confining potential energy is much larger than a possible gain energy with deformation. No bands are half-filled in Si or Mg wires, and thus the Peierls transition¹⁵ is not expected. The Peierls transition could be relevant if we were interested in wires that are half-filled, such as an alkali metal wire, but not here.

III. PHYSICAL PROCESSES BETWEEN WIRE AND ELECTRODES

For future atomic wire electronics, it is essential that technology be established to make a contact to electrodes that will be connected to batteries or to other atomic scale devices, and we need to clarify the physical processes of the contact at this stage. There will be particle exchange between the wire and the electrodes after contact, and if this occurs, the charge neutrality within the wire is not guaranteed, even in thermal equilibrium. (1) If the particle exchange through

the contact is extremely rare so that the electrons localize well inside the wire, the effective capacitance at the contact will be quite small and even the spin degeneracy will be lifted. The energy levels (renormalized to include the Coulomb interaction) are discrete, and the expectation value for the number of electrons in the wire is an integer. (2) If the contact allows some particle exchange, the electrons will not be strongly localized in the wire and the effective capacitance for this situation may be large. The spin degeneracy is not lifted, but the discrete nature of the allowed levels still prevails. The expectation value for the number of electrons is not necessarily an integer. (3) If the contact allows extremely frequent particle exchange, the electrons will spread over the entire system consisting of the electrodes and the wire, and the associated effective capacitance is huge, but the electron energy levels are no longer discrete. The particle exchange rate is determined by the physical processes around the contact, and is influenced by the separation between the atomic wire and the electrodes (barrier width), and the work function difference at the contact (barrier height). They can be independently tuned experimentally. As in macroscopic *p-n* junctions, it is expected that there are microscopic space-charge layers to absorb the work function difference between the wire and the electrodes.^{10,11} The widths are estimated to be on the order of several atomic layers (~ 1 nm) for the energy difference of $\sim 10^0$ eV according to a formula $w \sim (2\epsilon V / e N_e)^{1/2}$ with $N_e^{1/3} \sim 10^9 \text{ m}^{-1}$ and $\epsilon / \epsilon_0 \sim 10^1$.

We need to think of the finite-length effects of a wire. The band width W of the wire is determined by the lattice constant d and is a decreasing function of d until the bands cross.^{4,7} W does not depend very much on the number of atoms N as long as $N \gg 1$, and the average level separation is given by W/N .¹¹ These energy levels in principle correspond to the modes whose envelope functions are standing waves within the wire, and are well approximated with this picture. The highest and lowest states correspond to infrared and ultraviolet cutoff states in the wire structure, respectively. Typical $W(d)$ values are such that a Si metallic wire has a $W = 5$ eV for $d = 4 \text{ \AA}$, and a Mg semiconducting wire has a $W = 1.5$ (valence) and 2.7 eV (conduction) for $d = 5 \text{ \AA}$,⁴ so that the level separation is on the order of 10^{-1} eV for $N \sim 10^1$. If fully filled, the Fermi energy, ζ , lies in the middle of the highest-occupied and lowest-unoccupied levels. If partially filled, ζ lies at the highest occupied level. The electron filling can be changed by tuning the wire length and designing electrodes as shown below.

Figure 5(a) schematically shows the energy band alignment in equilibrium before contact, assuming that the electrode work function χ_e is greater than the wire work function χ_w . Energy lines for the opposite situation can be obtained with a vertical flip of the energy figures. ζ_e and ζ_w are the Fermi energies of the electrodes and the wire measured from the band bottom or the lowest level, respectively (both excluding the core levels). ϕ_e and ϕ_w are electrostatic energies for the electrodes and the wire, respectively. As is different from the usual macroscopic *p-n* junctions, the electrostatic

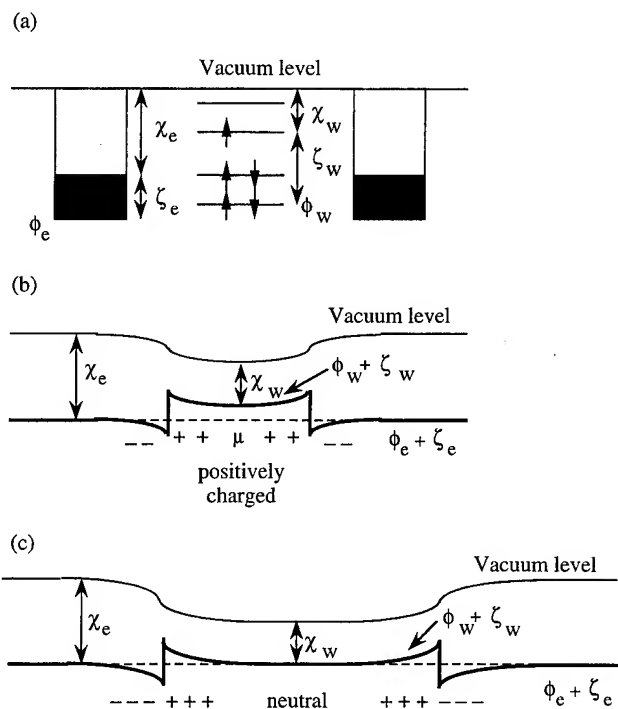


FIG. 5. Physical processes at contact between wire and electrodes: (a) Energy alignment before contact; (b) contact to short atomic wire; (c) contact to long atomic wire.

energy depends logarithmically on the distance from the center axis of the wire, and the energy lines should be understood as results of the appropriate averaging weighted with relevant electron wave functions.

After contact, electrons will spill out from the wire to the electrodes due to the work function difference $\chi_e > \chi_w$. Figures 5(b) and 5(c) indicate the equilibrium energy alignment for short and long wires, respectively. The thick lines are the summation of the electrostatic energy ϕ and the Fermi energy ζ , and the thin lines are vacuum levels. The electrochemical potential μ ,¹⁶ with a broken line, equals $\phi_e + \zeta_e$ deep inside the electrodes. Discrete energy levels (not shown) will be formed under the influence of ϕ , as in the case of an inversion channel with a triangular potential in a macroscopic heterojunction. If $\phi + \zeta - \mu$ is positive, that point is positively charged, and if negative, negatively charged. The discontinuity in electrostatic energy $(\phi_w + \zeta_w) - (\phi_e + \zeta_e)$ at the contact equals the work function difference $\chi_e - \chi_w$.^{10,16} If the wire is so short that the space charge layers at both ends touch as shown in Fig. 5(b), the charge neutrality is broken over the entire wire, and the electron filling for the discrete energy levels is no longer equivalent to that for an isolated wire. If the wire is long or the discontinuity in $\phi + \zeta$ is small so that the space-charge layers never touch, the charge neutrality holds with $\phi_w + \zeta_w = \mu$ in the middle as in Fig. 5(c). The electron filling in the middle is the same as that in an isolated wire, and this limit provides a situation of interacting Fermions described by the Tomonaga-Luttinger model.¹⁷

It has to be emphasized that there is certainly a case that

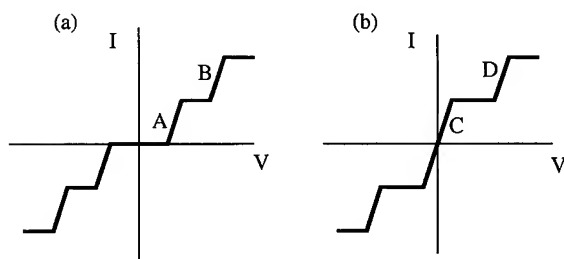


FIG. 6. Expected I - V characteristics with intermediate particle exchange at contact: (a) fully filled and (b) partially filled cases.

the electron filling in the wire with electrodes is completely different from that of the wire without electrodes as in Fig. 5(b). In an extreme situation, the electronic nature of the wire would be altered drastically, such that a metallic wire without electrodes will be insulating with electrodes or vice versa. In order to observe that a Si wire is metallic and a Mg wire is insulating, we need to make the wires long enough, much longer than the space charge layer width, or design the electrodes so that the work function difference will be minimized as in Fig. 5(c). The longer wires have their electronic identity while the shorter wires do not, and it is worth noticing this trivial, but important point.

IV. EXPECTED I - V CHARACTERISTICS WITH INTERMEDIATE PARTICLE EXCHANGE

With intermediate particle exchange, with $\chi_e \sim \chi_w$, we expect that the associated effective capacitance is large and the discreteness of energy levels is still relevant. Then, there will be two distinct patterns for metallic wire transport: the electrons are less likely to flow for a fully filled case and more likely for a partially filled case, where the maximum differential conductance is limited by the mode-selection process at the contact. The expected I - V patterns are schematically shown in Figs. 6(a) and 6(b) for fully and partially filled cases, respectively. For the fully filled case, the current will flow only when the applied voltage overcomes the separation of the discrete energy levels as in portion A, while for the partially filled case, the current will flow with an infinitesimal voltage as in portion C. The typical energy separation is on the order of 10^{-1} eV for an atomic wire with 10^1 atoms with a separation of 10^0 Å, and this gives a rough scale for the expected offset voltages. When a finite voltage is applied, the electrode Fermi energies in the source and drain split to support it. Every time the electrode Fermi energies cross a new wire level, it creates a sudden increase in current by forming steps in the I - V characteristics as shown in portions B and D. The current onsets have a finite gradient due to the mode-selection process at the contact. As is often pointed out in mesoscopy, such mode selection at the contact causes a finite conductance even though the transmission coefficient is unity, and we will observe the quantum unit of conductance (or its integer multiple).⁸ With this inherent effect, the differential conductance of the current onset cannot be infinite experimentally, but limited to be on the order of the quantum unit of conductance.

The current value at a step is determined by either the electron group velocity corresponding to the mode of interest or the transmission coefficient.^{18,19} Since one mode (a zero-dimensional mode where the momentum along the wire is also quantized) creates one discretized current independent of the voltage in the lowest order approximation, we will have stairlike I - V patterns. This stairlike shape is a general feature of the expected I - V characteristics and is independent of the details of the calculation model.

In the infinite wire limit, we can mathematically show that the quantum unit of conductance will be recovered, regardless of the boundary condition, which is a restatement of the usual cancellation of the electron state density and the electron group velocity.⁸ It has to be emphasized that this result is independent of the details of the dispersion or how the discrete energy levels have been formed in the original finite wire, and applies to quite general cases with a few modes, as is established in mesoscopy.⁸ The current is proportionate to the voltage in this limit simply because the number of modes participating in the transport is proportionate to the applied voltage, due to the continuous distribution of modes in the relevant energy region. This prediction is at least not contradictory to the experimental findings of the quantization of conductance in various metallic wire structures with atomic dimensions,²⁰ but not yet exactly verified experimentally. In fact, each studied geometry is not quite the one-dimensional atomic wires discussed here. Second, we cannot completely exclude the possibility that they have observed, e.g., bimodal transmission in the diffusive regime.

V. SUMMARY AND DISCUSSIONS

Various atomic structures with Si and Mg atoms are studied theoretically, and it is shown that Si wires and arrays are metallic, Mg wires are insulating, and Mg arrays have metallic and insulating phases. The calculation neglects the effects from the substrate, but they will not change the above results qualitatively.⁴ The electronic properties of wires cannot be discussed without specifying the electrodes, especially when the wire is short, because of the physical processes at the contact. The particle exchange rate at the contact determines the relative importance of the Coulomb interaction and the level broadening effect. With intermediate particle exchange along with $\chi_e \sim \chi_c$, depending on whether the highest occupied level is fully filled or partially filled, there would be differences in the I - V characteristics for metallic wire transport. An offset voltage is required for a current onset for the former, while an immediate current onset is expected for the latter, with a finite differential conductance on the order of the quantum unit of conductance. Each time the Fermi energies in the electrodes cross a new wire level with voltage, a step in the I - V characteristics results.

Lang has already discussed a special situation on how space charge layers form in Ref. 21 with a first-principle calculation. He studied the electronic properties of an adatom on a metallic surface, and assumed that the Fermi energy of the adatom aligned to that of the metallic surface. The adatom was either filled or depleted with electrons, depending

on the difference of the original work functions before contact. The discussion here is consistent with this if the adatom is regarded as the shortest wire in Fig. 5(b).

Lang studied the transport of an atomic wire with the first principle method in Ref. 22, where Al atomic wires with one or three atoms showed a larger conductance than that with two atoms. He assumed that the coupling of the wire atoms to the electrodes was so strong that the wire state densities were no longer discrete as in an isolated system, but had a significant width and overlap, showing some peaks only. This corresponds to a situation between the intermediate and extremely frequent particle exchange discussed above. Even in this case, there was a noticeable difference in transport properties between a two-atom wire, and one- and three-atom wires, and the current article is consistent with his result. In fact, if the Al wire is isolated, the parallel p state along the Al wire is modified more than the perpendicular p states because the $V_{pp\sigma}$ coupling is stronger than the $V_{pp\pi}$ coupling as discussed in Ref. 7. The highest occupied molecular level therefore originates from this parallel p state and is not degenerate, so that the highest occupied level is partially filled in one- and three-atom wires while it is fully filled in a two-atom wire. The former wires must have a better transport than the latter according to the view here, and his results showed it.

It is pointed out that there would be a Fermi energy pinning mechanism at the surface due to the surface states¹⁶ for a macroscopic junction, and the discontinuity in the electrostatic potential was not always a difference of two work functions. Such surface states do arise at the contact edge, but only when the lattice constant is so short such that the band crossing occurs, resulting in the Shockley's surface states.^{11,23} We need to avoid such short lattice constants in order to achieve mechanical wire stability, and both requirements are consistent. Alternatively, the surface contamination or the existence of foreign atoms on the surface would bring about these surface states, but they are unlikely under a ultrahigh vacuum STM environment, and the Fermi energy pinning through this route will not practically occur, either.

For electronics applications, there are some possible directions. One is to design devices making use of the Coulomb blockade effects with extremely small capacitances that can easily be obtained in the current atomic dimensions. Another is to create various junctions of different wires, where the Coulomb interaction will be suppressed by selecting a large-capacitance contact. This will lead to atomic wire electronics, to design materials with desired electronic properties by changing the atom arrangement, and employ the same device ideas in current semiconductor technology. For example, a Si wire is metallic and a Mg wire is insulating. Then, is the series junction of these two wires equivalent to a macroscopic metal-semiconductor junction showing rectifying characteristics? Does the junction structure of a metallic two-dimensional Mg array and a semiconducting Mg wire work as a metal-semiconductor junction? How can we dope an intrinsic semiconducting Mg wire, and how can we make

a p - n junction? Any possibility for the creation of light-emitting diodes? These points are left for future studies.

ACKNOWLEDGMENTS

The author is especially indebted to Professor W. A. Harrison for the discussions of band calculations and Professor M. Heiblum for the discussions of Coulomb interaction.

¹For example: D. M. Eigler and E. K. Schweizer, *Nature* (London) **344**, 524 (1990); J. A. Stroscio and D. M. Eigler, *Science* **254**, 1319 (1991); I.-W. Lyo and Ph. Avouris, *Science* **253**, 173 (1991); D. M. Eigler, C. P. Lutz, and W. E. Rudge, *Nature* (London) **352**, 600 (1991); H. J. Hamin, S. Chiang, H. Birk, P. H. Guethner, and D. Ruger, *J. Vac. Sci. Technol. B* **9**, 1398 (1991); M. F. Crommie, C. P. Lutz, and D. M. Eigler, *Science* **262**, 218 (1993); Ph. Ebert, M. G. Lagally, and K. Urban, *Phys. Rev. Lett.* **70**, 1437 (1993); H. Uchida, D. Huang, F. Grey, and M. Aono, *Phys. Rev. Lett.* **70**, 2040 (1993); Ph. Avouris, I.-W. Lyo, and Y. Hasegawa, *J. Vac. Sci. Technol. A* **11**, 1725 (1993); C. T. Salling and M. G. Lagally, *Surf. Sci.* **265**, 502 (1994).

²In the early time of band calculations, a gedanken experiment was often performed: plot the band widths as a function of lattice constant ranging from infinite to finite values, and clarify how bands originate from atomic levels, e.g., W. Shockley, *Electrons and Holes in Semiconductors* (Van Nostrand, Princeton, 1950) or J. C. Slater, *Quantum Theory of Matter*, (McGraw-Hill, New York, 1968). It has been recognized since then that the band structure can be modified by manipulation of the lattice constant. Due to the current development of atom manipulation technology, it is becoming possible to pursue this idea and proceed to atomic wire electronics.

³T. Yamada, Y. Takiguchi, D. Huang, and Y. Yamamoto, *Atomic Wire Circuit Network and Method*, US Patent 1995, pending; D. Huang and Y. Yamamoto, *Jpn. J. Appl. Phys.* **35**, 3734 (1996).

⁴T. Yamada, Y. Yamamoto, and W. A. Harrison, *J. Vac. Sci. Technol. B* **14**, 1243 (1996); T. Yamada and Y. Yamamoto, *Phys. Rev. B* **54**, 1902 (1996).

⁵For potential between closed-shell atoms, R. G. Gordon and Y. S. Kim, *J. Chem. Phys.* **56**, 3122 (1972).

⁶For extension to open-shell atoms, W. A. Harrison and R. Sokel, *J. Chem. Phys.* **65**, 379 (1976).

⁷W. A. Harrison, *Electronic Structure and Properties of Solids* (Freeman, San Francisco, 1980); *Surf. Sci.* **299/300**, 298 (1994); W. A. Harrison, *Phys. Rev. B* **24**, 5835 (1981); **31**, 2121 (1984).

⁸C. W. J. Beenakker and H. van Houten, in *Solid State Physics*, edited by Ehrenreich and D. Turnbull (Academic, San Diego, 1991), Vol. 44 and references therein.

⁹T. Yamada, *J. Vac. Sci. Technol. A* (to be published).

¹⁰S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981).

¹¹W. Shockley, *Electrons and Holes in Semiconductors* (Van Nostrand, Princeton, 1950).

¹²J. A. Stroscio, R. M. Feenstra, and A. P. Fein, *J. Vac. Sci. Technol. A* **5**, 838 (1987); J. A. Stroscio and R. M. Feenstra, in *Methods of Experimental Physics*, edited by J. A. Stroscio and W. J. Kaiser (Academic, San Diego, 1993), Vol. 27, Chap. 4.

¹³K. C. Pandey, *Phys. Rev. Lett.* **47**, 1913 (1981).

¹⁴There is another idea, which is to allow the foreign atoms to create chemical bonds with the substrate atoms by choosing an appropriate combination of the foreign atoms and substrate atoms with a suitable crystal orientation, requiring the electronic states to be localized within wires (or arrays) having an energy spectrum well inside the surface band gap of the substrate, S. Watanabe, Y. A. Ono, T. Hashizume, Y. Wada, J. Yamauchi, and M. Tsukada, *Phys. Rev. B* **52**, 10 768 (1996). It has to be emphasized that the band structure and the Fermi energy are generally different from those of an isolated wire. The wire and the substrate are unified electronically, and all the physical properties have to be designed for the entire system. For example, the creation of a second wire or the small fluctuation in substrate doping may alter the properties of the original wire completely. However, this route may open the door to higher temperature operation due to the excellent mechanical stability of the wire.

¹⁵C. Kittel, *Introduction to Solid State Physics*, 5th ed. (Wiley, New York, 1976).

- ¹⁶J. Bardeen, Phys. Rev. **71**, 717 (1947).
- ¹⁷S. Tomonaga, Prog. Theor. Phys. **5**, 544 (1950); J. M. Luttinger, J. Math. Phys. (N.Y.) **4**, 1154 (1963).
- ¹⁸For the group velocity picture, see H.-F. Chueng, Y. Gefen, E. K. Riedel, and W.-H. Shih, Phys. Rev. B **37**, 6050 (1988); and for the transmission picture, see Y. Imry, in *Directions in Condensed Matter Physics*, edited by G. Grinstein and G. Mazenko (World Scientific, Singapore, 1986) Vol. 1; R. Landauer, Z. Phys. B **68**, 217 (1987); M. Büttiker, Phys. Rev. Lett. **57**, 1761 (1986).
- ¹⁹How the group velocity and the transmission pictures are related to give a consistent, unified description for the current magnitude is left for future study. The general shape of the I-V characteristics is steplike, and a detailed model will just address how much the current and voltage steps are, respectively, in the limit of intermediate particle exchange. Those details are not of interest here.
- ²⁰U. Landman, W. D. Luedtke, N. A. Burnham, and R. J. Colton, Science **248**, 454 (1990); C. J. Muller, J. van Ruitenbeek, and L. J. de Jongh, Phys. Rev. Lett. **69**, 140 (1992); L. Olesen, E. Laegsgaard, I. Stensgaard, F. Besenbacher, J. Schiøtz, P. Stoltze, K. W. Jacobsen, and J. K. Nørskov, Phys. Rev. Lett. **72**, 2251 (1994); J. M. Krans, J. M. Van Ruitenbeek, V. V. Fisun, I. K. Yanson, and L. and J. De Jongh, Nature (London) **375**, 767 (1995); J. L. Costa-Krämer, N. García, P. García-Mochales, and P. A. Serena, Surf. Sci. **342**, L1144 (1995); J. I. Pascual, J. Méndez, J. Gómez-Herrero, A. B. Baró, N. Garcia, Uzi Landman, W. D. Luedtke, E. N. Bogachek, and H.-P. Cheng, Science **267**, 1793 (1995).
- ²¹N. D. Lang, Phys. Rev. Lett. **56**, 1164 (1986).
- ²²N. D. Lang, Phys. Rev. B **52**, 5335 (1995).
- ²³W. Shockley, Phys. Rev. **56**, 317 (1939).

Morphological and compositional variations in strain-compensated InGaAsP/InGaP superlattices

R. S. Goldman^{a)}

Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213 and Department of Materials Science and Engineering, The University of Michigan, Ann Arbor, Michigan 48109-2136

R. M. Feenstra^{b)}

Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

C. Silfvenius, B. Stålnacke, and G. Landgren

Department of Electronics, Royal Institute of Technology, Kista, Sweden

(Received 12 February 1997; accepted 14 April 1997)

We have investigated the properties of strain-compensated InGaAsP/InGaP superlattices, grown by metalorganic vapor phase epitaxy, with and without InP interlayers inserted in the InGaP barrier. Using cross-sectional scanning tunneling microscopy and spectroscopy, we observe lateral variations in layer thickness and electronic properties. When the number of superlattice periods is increased from 8 to 16, the growth front develops large undulations in the top two to four superlattice periods. For structures with InP layers inserted in the InGaP barrier, only slight undulations of the top superlattice periods occur. We discuss the origins of the growth front undulations in terms of the elastic relaxation of strain arising from thickness and/or composition variations in the superlattice layers. Finally, we observe a fourfold periodicity of the [001] atomic spacing, presumably arising from atomic ordering in the alloys. © 1997 American Vacuum Society. [S0734-211X(97)06604-3]

I. INTRODUCTION

InGaAsP-based multiple quantum well structures with alternating compressive- and tensile-strained layers are promising for optoelectronic device applications, such as light sources and detectors in optical fiber communications systems.¹ Although the structures are intended to be strain balanced, the optical properties are often degraded by residual-strain related effects.² In particular, lateral variations in surface morphology and/or alloy composition have been observed.^{3,4} The interplay between these phenomena, and their effects on optical properties are not well understood. A detailed study of the nanometer-scale structure and electronic properties is essential for understanding the residual-strain related effects and their impact on photoluminescence (PL) efficiency.

In this article, we present cross-sectional scanning tunneling microscopy (STM) and spectroscopy studies of a series of InGaAsP/InGaP superlattices produced by metalorganic vapor phase epitaxy (MOVPE). The structures have an effective bandgap near 1.3 μm , and have a constant In/Ga ratio in both the well and barrier. Apart from fabrication convenience, this results in very deep electron wells, thereby minimizing carrier leakage which leads to poor high temperature performance in many 1.3 μm laser designs. In some of the superlattices we find lateral variations in the effective band gaps and layer thicknesses of the ternary and quaternary alloy layers. In addition, we observe growth front undulations whose amplitude depends on the total number of superlattice periods and the presence of InP interlayers in the InGaP barrier. Furthermore, the PL efficiency is reduced for structures

with significant growth front undulations. In some cases, we find a fourfold periodicity of the (001) planes, presumed to arise from atomic ordering of the alloys. We discuss the origins of the growth front undulations in the context of the elastic relaxation of strain arising from thickness and/or composition variations in the superlattices.

The article is organized as follows. Section II describes the experimental details, including growth conditions and characterization by x-ray diffraction and PL. STM images and spectroscopy data, as well as an illustration of the structural degradation process, are presented in Sec. III. Finally, conclusions are given in Sec. IV.

II. EXPERIMENTAL DETAILS

The samples were prepared by MOVPE at a pressure of 50 mbar, using trimethylindium, trimethylgallium, AsH₃, and PH₃ as source gases. The undoped structures were grown on sulfur-doped ($N_d = 5 \times 10^{18} \text{ cm}^{-3}$) InP substrates, at a substrate temperature of 680 °C.² The targeted structures included 8 or 16 period superlattices (multiquantum wells) surrounded by a 240 nm InP buffer layer and 150 nm InP cap. Each superlattice period consisted of a quantum well of 5 nm In_{0.89}Ga_{0.11}As_{0.55}P_{0.45} and a barrier of either 9 nm In_{0.89}Ga_{0.11}P or 4.5 nm In_{0.89}Ga_{0.11}P+10 nm InP+4.5 nm In_{0.89}Ga_{0.11}P. Since the quantum well and barrier have 1.0% compressive and 0.8% tensile lattice-mismatch strains, relative to the InP substrate, the strain is nearly balanced. In principle, it should be possible to grow multiple superlattice periods, without degradation of the optical properties. However, the measured 300 K PL intensities were highest and lowest for the 8- and 16-period superlattices, respectively, with intermediate values for the 16-period superlattices with

^{a)}Electronic mail: rsgold@engin.umich.edu

^{b)}Electronic mail: feenstra@andrew.cmu.edu

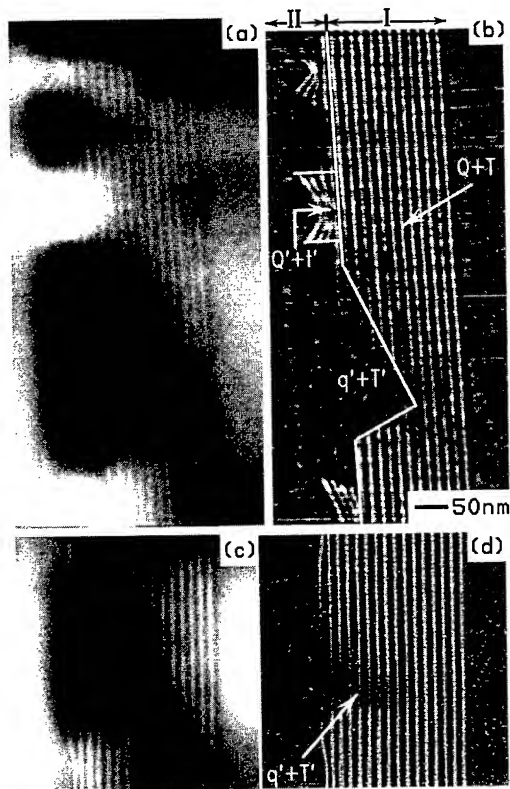


FIG. 1. Large-scale topographic images of 16-period InGaAsP/InGaP superlattices, acquired at sample bias voltages of (a),(b) +2.5 V and (c),(d) -2.0 V. The InGaAsP (quaternary) and InGaP (ternary) layers appear bright and dark, respectively. The grey-scale ranges displayed are (a) 15 Å, (b) 1.7 Å, (c) 10 Å, and (d) 0.7 Å, where a local background subtraction was performed in (b) and (d).

InP interlayers.² High-resolution x-ray diffraction spectra showed significantly broadened satellite peaks for the 16-period superlattices in comparison with the 8-period superlattices, indicating a less well defined periodic structure in the 16-period superlattices in comparison with the 8-period superlattices.^{5,6} For the 16-period superlattices with InP interlayers in the InGaP barrier, the satellite peaks in the x-ray diffraction spectrum were well defined, but direct comparison with the other two spectra is not straightforward. Although the x-ray results indicate a link between a less pronounced periodic structure and reduced PL efficiency, they do not provide detailed information about how the superlattice periodicity has been reduced.

For STM, the samples were cleaved to expose a $(\bar{1}10)$ surface,⁷ in an ultrahigh vacuum chamber with base pressure $<5 \times 10^{-11}$ Torr. STM was performed with both electrochemically etched single crystal $\langle 111 \rangle$ -oriented W and commercially available Pt-Ir tips. The tips were cleaned by *in situ* electron bombardment and characterized by *in situ* field emission microscopy. Images were obtained with a constant tunnel current of 0.1 nA, and sample bias voltages as described below. Details of the STM design,⁸ cleavage procedure,⁹ and spectroscopic methods¹⁰ have been presented elsewhere.

III. RESULTS AND DISCUSSION

Figures 1(a)–1(d) show large scale STM topographic images of the 16-period InGaAsP/InGaP superlattices, displayed with the growth direction from right to left. The empty and filled state images presented in Figs. 1(a) and 1(c), respectively, were acquired at different locations on the sample. In both Figs. 1(a) and 1(c), the InGaAsP (quaternary) and InGaP (ternary) layers appear as alternating bright and dark layers sandwiched between thick layers of InP (the buffer and cap). Since the InGaAsP/InGaP system is a type I superlattice, both electrons and holes are confined in the InGaAsP layers. Thus, for a given tip-sample bias voltage in the STM, a greater number of filled or empty states contribute to the tunnel current for the InGaAsP layers compared to the InGaP layers. In order to maintain a constant tunneling current, the tip moves towards the (cleaved) surface in the vicinity of the InGaP, and away from the (cleaved) surface in the vicinity of the InGaAsP. As a consequence, the InGaP and InGaAsP layers appear as dark and bright layers, respectively.

The images of Fig. 1 clearly reveal structural degradation of the superlattice layers. In particular, large growth front undulations and large lateral variations in image contrast are apparent. The (001) growth front undulations arise from relaxation of strain accumulated in the superlattices during growth. The lateral variations in image contrast on the $(\bar{1}10)$ cleavage surface are due to a combination of band edge shifts due to varying thicknesses, alloy compositions, and strains in the layers, as well as a distortion of the cleavage surface associated with relaxation of strain. These effects will be discussed in more detail below.

Figures 1(b) and 1(d) show topographic images in which the local background (formed by averaging the data over an area of $20 \times 20 \text{ nm}^2$) has been subtracted from Figs. 1(a) and 1(c), respectively. In Figs. 1(b) and 1(d), much of the lateral image contrast variation is suppressed, such that lateral variations in the layer thicknesses are evident. The observed structural degradation process is divided into two regimes; phase I—where the superlattice period is conserved, and phase II—where the superlattice period varies and large undulations of the surface occur. In the phase I regime, significant changes in layer thicknesses are observed, as seen most clearly in filled state images such as Fig. 1(d), with thin quaternary layers and thick ternary layers (the sum of their thicknesses is fixed). Closest to the substrate is region $Q + T$, where the superlattice layer thicknesses are fixed at the targeted values. In Fig. 2(a), we present a high-resolution view of region $Q + T$, acquired at a sample bias voltage of +2.5 V. Fringes with a spacing of 5.8 and 5.9 Å, corresponding to the (001) planes of the ternary and quaternary alloys, are observed in the dark and bright layers, respectively. The image consists primarily of three superlattice periods, with typical ternary and quaternary thicknesses 10.0 and 4.7 nm, close to the targeted values.

Additional regions denoted in Figs. 1(b) and 1(d) are $q' + T'$, with thin quaternary and thick ternary layers, and $Q' + t'$, with thick quaternary and thin ternary layers. In both

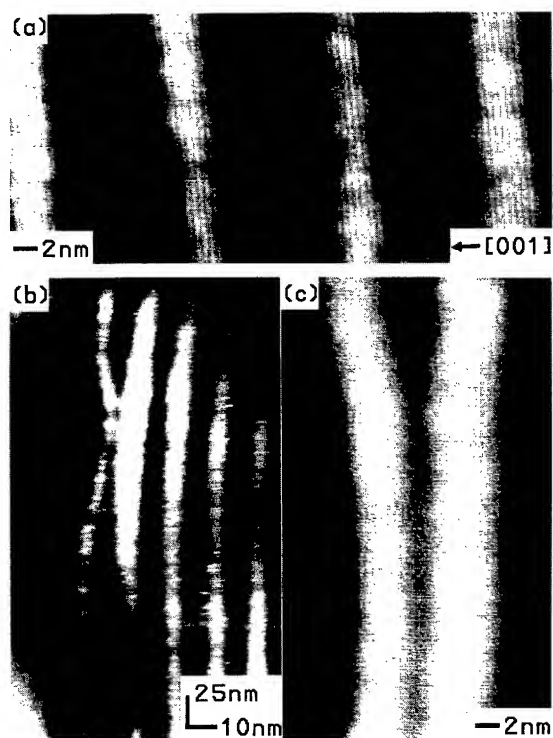


FIG. 2. Close-up views of 16-period InGaAsP/InGaP superlattices. The high-resolution image of region $Q+T$ displayed in (a), was acquired at a sample voltage of +2.5 V. The images in (b) and (c), acquired at sample voltages of -1.8 and +1.8 V, show a close-up view of an undulation, and a high-resolution view of region $Q'+t'$, respectively. The vertical scale in (b) is compressed by a factor of 2.5, as indicated on the image.

cases, the primes indicate a possible change in alloy composition in the layers. The thickness variations are most evident in the close-up views shown in Figs. 2(b) and 2(c), acquired at sample bias voltages -1.8 and +1.8 V, respectively. Figure 2(b) displays a close-up view of a (growth front) morphological undulation, where the top five superlattice periods in regions $q'+T'$ and $Q'+t'$ (phase II) are displayed. Lateral variations in the ternary and quaternary layer thicknesses are evident. In Fig. 2(c), we show a high-resolution view of region $Q'+t'$ (also phase II). Similar to Fig. 2(a), fringes corresponding to the (001) planes of the ternary and quaternary alloys are observed. The image consists of two superlattice periods, with ternary layers as thin as 1.2 nm. It is evident from this image that the (001) lattice planes remain aligned along the [110] direction while the profiles of the ternary and quaternary layers do not. This indicates a change in surface normal during the growth of the undulations (i.e., faceting), similar to that reported in InGaP/InAsP structures.⁵ We observe angles between the (001) plane normal and the normal to the layer profile ranging from 20° to 40°, corresponding to (114), (113), or (112) facets.

Some of the accumulated strain in the 16-period superlattices is probably relaxed elastically by the formation of the observed growth front undulations. In addition, in some cases, the elastic relaxation is accompanied by plastic relaxation, as we see in the image in Fig. 3. In this empty state image, acquired at a sample bias voltage of +2.5 V, a sur-

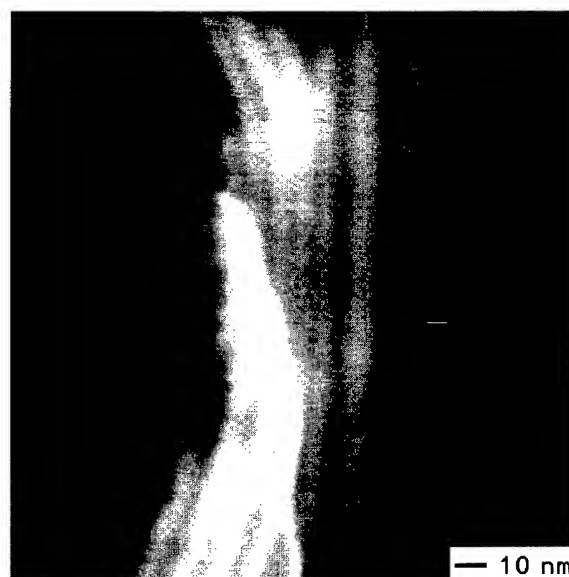


FIG. 3. Topographic image acquired near a valley of an undulation, at a sample bias voltage of +2.5 V. A surface step extending out from the cusp in the valley is observed, and attributed to a dislocation.

face step extending out from the cusp in the valley is observed. We attribute this step to a screw-type dislocation occurring near the cusp of the valley. We cannot say at present whether the dislocation was nucleated before the valley formed (i.e., due to high stress at the cusp), or alternatively, that the dislocation formed first and its associated stress relaxation produced a large variation in growth rates which then led to the valley formation. In either case, the existence of the dislocation is a clear indication of the high stress levels in the region near the cusp of the valley.

In order to quantify the lateral band gap variations observed in bias-dependent topography images, we performed spatially resolved spectroscopy measurements in the different regions outlined in Figs. 1(b) and 1(d). In Fig. 4, the normalized conductance versus sample bias voltage is plotted for (a) the quaternary layers, (b) the ternary layers, and (c) the InP substrate or cap layers. For each layer, the normalized conductance versus sample bias voltage was collected in the $Q+T$, $Q'+t'$, and $q'+T'$ regions (the InP spectra were acquired *nearby* those labeled regions). All the spectra display well defined band edges, as marked by dashed lines in Fig. 4. The nonzero conductance observed within the gap is the "dopant-induced" component, which arises from electrons tunneling out of filled conduction band states.¹⁰ Table I summarizes the results of the band gap measurements. The good agreement between theory and experiment for the InP spectra indicates that tip-induced band bending is not significant in these measurements.¹⁰ As mentioned earlier, the InGaAsP/InGaP system is a type I superlattice, with both electrons and holes confined in the InGaAsP (quaternary). Therefore, the effective band gap of the ternary corresponds to the energy difference between the InGaP conduction band edge and InGaP light hole valence

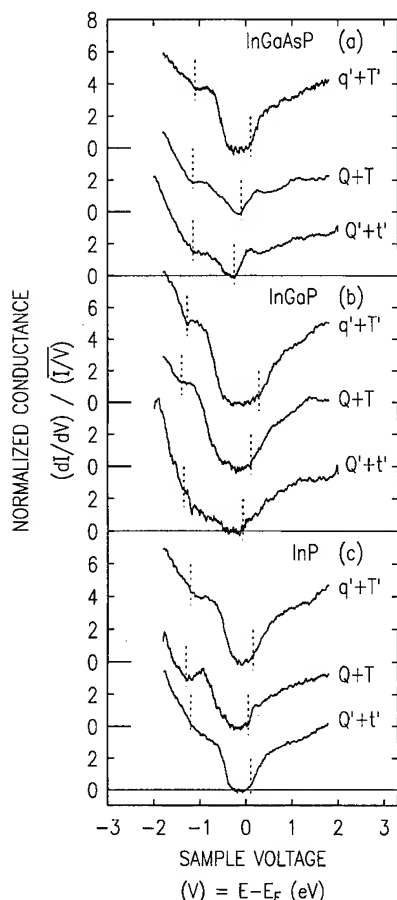


FIG. 4. Spatially resolved spectroscopy results, acquired on the (a) quaternary layers, (b) ternary layers, and (c) on the InP substrate or cap layers. In each case, the location of the spectra in terms of the $Q+T$, $Q'+t'$, or $q'+T'$ regions is indicated (the InP spectra were acquired *nearby* those labeled regions).

band, while the effective band gap of the quaternary corresponds to the energy difference between the first electron (e_0) and first hole (h_0) subbands ($E_{e_0-h_0}$) in the InGaAsP.

We have calculated the expected effective band gaps of the ternary and quaternary layers of the superlattice, taking into account the lattice mismatch strain, and possible varia-

tions in layer thickness and/or alloy compositions. Strain effects were included assuming biaxial strain and using the appropriate deformation potentials,¹¹ and subband energies were computed using the transfer matrix method.¹² The results are shown in Table I. For region $Q+T$, with ternary and quaternary alloy compositions and layer thickness fixed at the targeted values, the calculated effective band gaps of the ternary and quaternary are 1.41 and 0.94 eV, respectively. These values are within the error of our measured values, 1.50 ± 0.10 and 1.05 ± 0.10 eV for the ternary and quaternary, respectively. In region $q'+T'$, with thick ternary and thin quaternary layers, the calculated effective band gaps of the ternary and quaternary are 1.41 and 1.13 eV, respectively. In this case, the calculated quaternary band gap is within the error measured value, 1.20 ± 0.10 eV, but the calculated ternary band gap is smaller than our measured value, 1.56 ± 0.10 eV. Thus, additional factors, such as composition variations and/or lateral strain variations (nonbiaxial strain) may contribute to the larger measured effective gap. If we consider a composition variation of +5% Ga ($\text{In}_{0.84}\text{Ga}_{0.16}\text{P}$), the calculated effective band gap of the ternary is slightly increased to 1.44 eV, which is closer to our measured value discussed above. Additional contributions due to nonbiaxial strain will be discussed below. The calculated effective band gaps for the ternary and quaternary in region $Q'+t'$ are 1.41 and 0.88 eV, respectively. Similar to region $q'+T'$, the calculated effective gap of the quaternary is within the error of our measured value, 0.90 ± 0.10 eV, but that of the ternary is larger than our measured value, 1.28 ± 0.10 eV. If we consider a composition variation of -5% Ga ($\text{In}_{0.94}\text{Ga}_{0.06}\text{P}$), the calculated effective band gap of the ternary is decreased to 1.38 eV, which is within the error of the measured value.

Considering now the effects of lateral strain variations (nonbiaxial strain) on the observed band gaps, it is clear that in the $q'+T'$ region there will be additional tensile strain due to the larger amount of ternary (small lattice constant) material. If we approximate the $q'+T'$ material as being composed *entirely* of ternary material T' , with mismatch $f = (a_{\text{InP}} - a_{\text{InGaP}})/a_{\text{InGaP}} = +0.0079$, and further consider this

TABLE I. Comparison of measured and computed band gaps, ignoring corrections due to lateral variations in strain. Thickness of the quaternary layers are 5, 2, and 10 nm for the $Q+T$, $q'+T'$, and $Q'+t'$ regions, respectively. Layers are coherently strained to InP. All results are in eV.

Material	$Q+T$	$q'+T'$	$Q'+t'$
Measured band gaps			
InP	1.35 ± 0.10	1.35 ± 0.10	1.30 ± 0.10
$\text{In}_{0.89}\text{Ga}_{0.11}\text{P}$	1.50 ± 0.10	1.56 ± 0.10	1.28 ± 0.10
$\text{In}_{0.89}\text{Ga}_{0.11}\text{As}_{0.55}\text{P}_{0.45}$	1.05 ± 0.10	1.20 ± 0.10	0.90 ± 0.10
Calculated effective band gaps			
InP	1.35	1.35	1.35
$\text{In}_{0.89}\text{Ga}_{0.11}\text{P}$	1.41	1.41	1.41
$\text{In}_{0.89}\text{Ga}_{0.11}\text{As}_{0.55}\text{P}_{0.45}$	0.94	1.13	0.88
$\text{In}_{0.84}\text{Ga}_{0.16}\text{P}$		1.44	1.44
$\text{In}_{0.94}\text{Ga}_{0.06}\text{P}$		1.38	1.38

region to be a spherical inclusion embedded in InP, then the strain in the inclusion can be computed using an analytic solution for the elasticity equations for this geometry.¹³ Computing band edge shifts using the appropriate deformation potentials,¹¹ we find that the band gap of the T' region will be reduced by 0.050 eV compared with bulk $\text{In}_{0.89}\text{Ga}_{0.11}\text{P}$. This value must be compared with the expected gap of the ternary material in the $Q+T$ region, which is under biaxial strain leading to a band gap reduction of 0.101 eV. Thus, the net effect is an *increase* of the ternary gap in the $q'+T'$ material of about 0.05 eV compared to that in $Q+T$, which leads to better agreement between the theory and observed gap energies. Considering the thin quaternary layers in the $q'+T'$ regions to be biaxially strained onto the T' material leads to only an 0.01 eV increase in their band gap compared to the $Q+T$ material. Finally, for the $Q'+t'$ region, the strain is somewhat uncertain due to observed dislocation formation, so we do not attempt any further corrections to those band gap estimates.

In summary, the data indicate 0.1–0.2 eV variations in the effective band gaps of both the ternary and quaternary alloys, in the $Q+T$, $Q'+t'$, and $q'+T'$ regions. For imaging voltages with magnitude ≥ 2 V, this amount of band edge shift would lead to variations in tip-sample spacing of ≤ 1 Å, which is much less than the 10 Å contrast variation observed in the constant-current images of Figs. 1(a) and 1(c). Thus, additional topographic (nonelectronic) factors must be contributing to the observed image contrast. A topographic effect to consider is the influence of strain relaxation of the cleavage surface itself, as has been previously suggested by Pinnington *et al.* in their study of Si/Ge heterostructures.¹⁴ In our case, as discussed above, the $q'+T'$ layers have additional tensile strain compared to the $Q+T$ region, and similarly the $Q'+t'$ region is expected to have additional compressive strain. Thus, when the sample is cleaved, these regions will relax inwards and outwards from the cleavage face, respectively, in agreement with the observed sign of the contrast in Fig. 1. Since this surface distortion is a purely topographic effect, we do not expect it to significantly influence our measurements of effective band gap discussed above.

When InP interlayers are inserted in the InGaP barrier, the structural degradation of the growth front is greatly reduced. In Figs. 5(a) and 5(b), we present large-area topographic and conductance (dI/dV) images of 16-period superlattices with InP interlayers inserted in the InGaP barrier. The images are displayed with the growth direction from right to left. In the empty state topographic image presented in Fig. 5(a), acquired at a sample bias voltage of +2.5 V, a sequence of dark, bright, dark, and medium layers, corresponding to the ternary, quaternary, ternary, and InP, is sandwiched between an InP buffer and cap layers. The image also reveals lateral contrast variations similar to the superlattices without InP interlayers [as in Figs. 1(a) and 1(c)]. Thickness variations in the ternary and quaternary layers are most evident in the

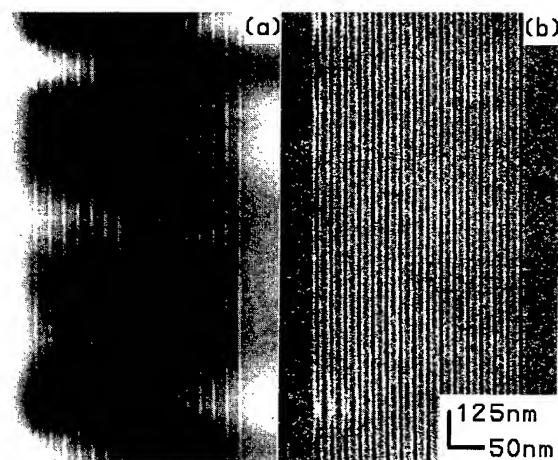


FIG. 5. Large-area (a) topography and (b) conductance images of 16-period InGaAsP/InGaP superlattices, with InP interlayers inserted in the InGaP barriers, acquired at a sample bias voltage of +2.5 V. The vertical scale in the images is compressed by a factor of 2.5, as indicated. The grey-scale range displayed in (a) is 6.0 Å.

conductance image shown in Fig. 5(b). In this case, only slight undulations of the top superlattice period are observed, in agreement with related studies.⁶

We have shown above several images illustrating lateral variations in layer thickness, effective band gap, and surface morphology in InGaAsP/InGaP superlattices. Here, we present a model relating the degradation of the structures with the buildup of strain in the superlattices. In phase I, shown in Fig. 6(a), we observe that the process begins with the formation of a thin region of a quaternary layer. We believe this nucleation event is random, since no extrinsic nucleation sources have been observed. Deposition of the next ternary layer will produce a strain variation at the surface due to the quaternary layer thickness variation.^{15,16} Assuming that the growth rate depends on strain, the resulting more negative (compressive) in-plane strain would favor the attachment of small atoms.¹⁷ In this case, the strain variation will inhibit further quaternary growth and favor ternary growth. Thus, a region with thin quaternary and thick ternary layers ($q'+T'$) is formed. We observe that this region propagates and grows in size laterally as more layers are deposited. Alloy composition variations may also occur, with Ga- and/or P-rich $q'+T'$ regions (we cannot distinguish between a random variation in thickness or composition which acts to nucleate the degradation process). Similar effects have been observed in InGaP/InAsP (Ref. 3) and Si/Ge (Ref. 18) superlattices. As the process shown in Fig. 6(a) continues, the lateral inhomogeneities in layer thickness and possible compositional variations will increase, leading to large lateral variations in the strain of the layers. This accumulated lateral strain then begins to relax elastically and/or plastically in phase II.

Phase II of the degradation process is characterized by a superlattice period which is no longer fixed. Deep valleys form in the morphology. The resulting growth front undulations may act to relieve stress (elastic relaxation). In addi-

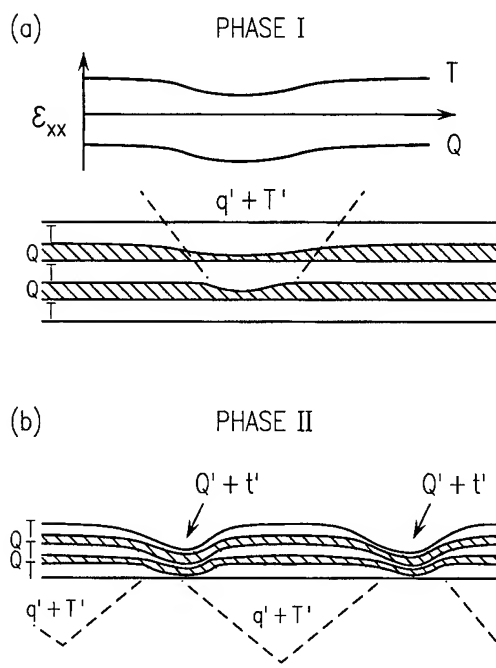


FIG. 6. Illustration of the buildup of strain in the degraded portions of the superlattice. In (a), the superlattice period is conserved (phase I). A quaternary layer Q which has a thin region is randomly formed. Deposition of the next ternary layer T will produce strain variations at the surface, due to the thickness variation of the quaternary layer. This region, denoted $q' + T'$, propagates and grows in size laterally as more layers are deposited. As this process continues, lateral variations in strain will accumulate and at some point begin to relax. As shown in (b), the superlattice period then varies and large undulations of the surface develop (phase II), arising from a combination of elastic and plastic strain relaxation. Regions with thin ternary layers, $Q' + t'$, are formed.

tion, dislocations are observed to form near the cusps of the valleys (plastic relaxation). The $Q' + t'$ regions forming near the valleys have very thin ternary layers. As in phase I, we associate this change in layer thickness with a reduced growth rate, due in this case to additional tensile in-plane strain in this region. This strain may also induce alloy compositional variations in this region.

For the superlattices grown with InP interlayers, the propagation of strain buildup in the structure is reduced. This may be due in part to the lack of alloy decomposition in a binary alloy. As a result, the propagation of alloy composition variations (and the resultant strain variations) through the structure is essentially filtered by the InP interlayers. The nonzero thickness of the InP layers leads to a reduction in strain variations on the upper interface of each InP layer compared to the lower interface. Also, it should be noted that since the average strain in the InP layers is zero (the substrate is InP), and the strain dependence of the chemical potential varies linearly with average strain,¹⁹ thickness variations in the InP layers are neither expected nor observed. The superlattices with InP interlayers have remained in phase I of the degradation process, and the size of the thickness and/or composition modulated regions is limited. Based on our observations, it would appear to be possible to grow many

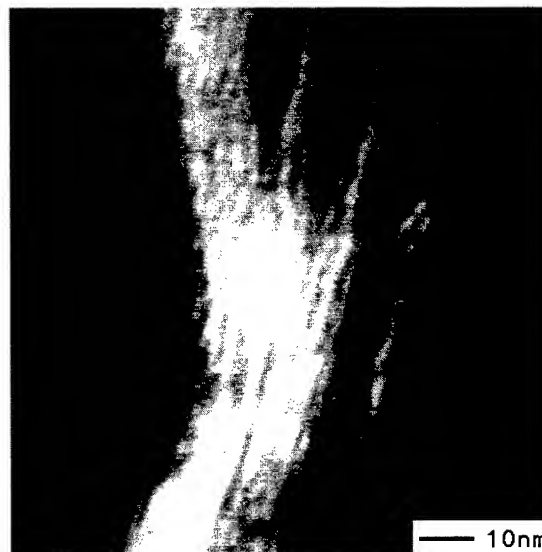


FIG. 7. Topographic image of the top six to seven superlattice periods, in the vicinity of a valley and nearby hill of an undulation, acquired at a sample bias voltage of +2.5 V. The grey-scale range displayed is 3.2 Å. Fringes with a spacing of 24 Å, corresponding to a fourfold multiple of the (001) atomic spacing, are observed.

more periods of superlattices with InP interlayers without severe (phase II) degradation of the structure.

In addition to the lateral variations in layer thickness, effective band gap, and growth front morphology, in some regions we observe a multiplicity in the atomic periodicity, as displayed in Fig. 7. That image shows the top six to seven superlattice periods, in the vicinity of a valley and nearby hill of an undulation, acquired at a sample bias voltage of +2.5 V. Lateral variations in layer thicknesses are observed, as discussed in Figs. 1–3. In addition, fringes with a spacing of 24 Å, are evident. Similar to the lattice planes in Fig. 2, the fringes remained aligned along the [110] direction while the profiles of the ternary and quaternary layers do not. This periodicity corresponds to a fourfold multiple of the [001] atomic spacing, suggesting atomic ordering in the alloys [this fourfold spacing is observed on a (110) plane, so it would be consistent, e.g., with a fourfold spacing of (111) planes]. Although atomic ordering has been previously discussed in compound semiconductor alloys,²⁰ this fourfold periodicity has not been previously reported. Furthermore, it is not explained by either CuPt or CuAu ordering, which would lead to a twofold multiple of the (001) atomic spacing.^{20,21} Remarkably, we observe that this fourfold periodicity extends into the overlying InP layers, as seen on the left-hand side of Fig. 7. We suggest that this apparent fourfold periodicity in the pure InP arises simply from an extension of the charge density perturbations that exist in the alloy due to the Brillouin zone folding that accompanies the alloy ordering. This charge density perturbation may be enhanced by relaxation of the (110) surface. The amplitude of this perturbation in the InP decays to zero at a distance of about 50 nm from the InGaP/InGaAsP superlattice. Further work is required to understand the origin of the fourfold periodicity and its impli-

cations on the present models for ordering in compound semiconductors.

IV. CONCLUSIONS

We have investigated the properties of InGaAsP/InGaP superlattices, grown with or without InP layers inserted in the InGaP barriers. Spatially resolved spectroscopy reveals variations in the effective band gaps of the superlattice layers. STM topography images indicate lateral variations in the layer thickness of the ternary and quaternary alloys of the superlattices. In some cases, we find a fourfold periodicity of the (001) planes, presumed to arise from atomic ordering in the alloys. When the number of superlattice periods is increased from 8 to 16, the growth front develops large undulations involving the top few superlattice periods. This effect is reduced in structures with InP layers inserted in the InGaP barrier. Furthermore, the PL efficiency decreases for structures with significant growth front undulations,² due to the inhomogeneous distribution of quantum wells and/or formation of dislocations in the structure. A model is presented describing the formation of the surface undulations in terms of accumulated strain due to the propagation of layer thickness and/or composition variations. Similar effects are expected to occur in other strained-layer materials systems.

ACKNOWLEDGMENTS

The work at Carnegie Mellon was supported in part by National Science Foundation Grant No. DMR-9615647. Discussions with S. Mahajan, M. Skowronski, J. Tersoff, and T. Tiedje are gratefully acknowledged.

- ¹E. P. O'Reilly and A. R. Adams, *IEEE J. Quantum Electron.* **30**, 366 (1994).
- ²C. Silfvenius, B. Stålnacke, and G. Landgren, *J. Cryst. Growth* **170**, 122 (1997).
- ³A. Ponchet, A. Rocher, A. Ougazzaden, and A. Mircea, *J. Appl. Phys.* **75**, 7881 (1994).
- ⁴G. Patriarche, A. Ougazzaden, and F. Glas, *Appl. Phys. Lett.* **69**, 2279 (1996).
- ⁵M. A. G. Halliwell, M. H. Lyons, S. T. Davey, M. Hockly, C. G. Tuppen, and C. J. Gibbings, *Semicond. Sci. Technol.* **4**, 10 (1989).
- ⁶C. R. Wie, *Mater. Sci. Eng. R. Rep.* **13**, 1 (1994).
- ⁷We define the growth directions as (001), and identify the $\langle 110 \rangle$ directions by anisotropic etching with HCl. In general, we find that cleaving the (110) face most often produces atomically flat surfaces. In the present study, we have not attempted cleaves of the (110) face.
- ⁸R. M. Feenstra, *Proceedings of the 21st International Conference on the Physics of Semiconductors*, edited by P. Jiang and H.-Z. Zheng (World Scientific, Singapore, 1992), p. 357.
- ⁹R. M. Feenstra, E. T. Yu, J. M. Woodall, P. D. Kirchner, C. L. Lin, and G. D. Pettit, *J. Vac. Sci. Technol. B* **61**, 795 (1992).
- ¹⁰P. Mårtensson and R. M. Feenstra, *Phys. Rev. B* **39**, 7744 (1988); R. M. Feenstra, *Phys. Rev. B* **50**, 4561 (1994).
- ¹¹C. G. Van de Walle, *Phys. Rev. B* **39**, 1871 (1989).
- ¹²B. Jonsson and S. T. Eng, *J. Quantum Electron.* **26**, 2025 (1990).
- ¹³R. S. Goldman, R. M. Feenstra, B. G. Briner, M. L. O'Steen, and R. J. Hauenstein, *J. Electron. Mater.* (in press).
- ¹⁴T. Pinnington, A. Sanderson, T. Tiedje, T. P. Pearsall, E. Kasper, and H. Presting, *Thin Solid Films* **222**, 259 (1992).
- ¹⁵R. M. Feenstra and M. A. Lutz, *J. Appl. Phys.* **78**, 6091 (1995).
- ¹⁶J. Tersoff, C. Teichert, and M. G. Lagally, *Phys. Rev. Lett.* **76**, 1675 (1996).
- ¹⁷J. E. Guyer and P. W. Voorhees, *Phys. Rev. B* **54**, 1170 (1996).
- ¹⁸E. Carlino, L. Tapfer, and H. von Känel, *Appl. Phys. Lett.* **69**, 2546 (1996).
- ¹⁹T. Pinnington, C. Lavoie, T. Tiedje, B. Haveman, and E. Nodwell, *J. Vac. Sci. Technol. B*, these proceedings.
- ²⁰See, for example, A. Zunger and S. Mahajan, in *Handbook on Semiconductors*, edited by S. Mahajan (North-Holland, Amsterdam, 1994), Vol. 3, p. 1399, and references therein.
- ²¹S. Froyen and A. Zunger, *Phys. Rev. B* **53**, 4570 (1996).

Spatial variations in luminescence and carrier relaxation in molecular beam epitaxial grown $(\text{InP})_2/(\text{GaP})_2$ quantum wires

Y. Tang and D. H. Rich^{a)}

*Department of Materials Science and Engineering, Photonic Materials and Devices Laboratory,
University of Southern California, Los Angeles, California 90089-0241*

A. M. Moy and K. Y. Cheng

*Department of Electrical and Computer Engineering, Center for Compound Semiconductors
Microelectronics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801*

(Received 14 January 1997; accepted 16 April 1997)

Phase separation in III–V semiconductors has led to a unique method for fabricating quantum wires via a strain induced lateral ordering process. Quantum wire (QWR) arrays were formed during the gas source molecular beam epitaxial (MBE) growth of $(\text{InP})_2/(\text{GaP})_2$ bilayer superlattices (BSLs) and were studied by time-resolved and linearly polarized cathodoluminescence. Nonlinear optical properties, such as phase-space filling effects, were observed to be indicative of the QWR nature of the samples. Samples prepared by gas source MBE were found to have a greater uniformity, smaller QWRs, and higher optical quality in comparison to those obtained by metal–organic chemical vapor deposition. Misfit dislocations also formed in one of the BSL samples, indicating a partial strain relaxation at the GaAs/InGaP and BSL/InGaP interfaces. The carrier relaxation, transport, and collection in the QWRs were studied with time-resolved cathodoluminescence. © 1997 American Vacuum Society. [S0734-211X(97)07504-5]

I. INTRODUCTION

The majority of III–V alloys exhibit miscibility gaps at low temperature and this can result in a lateral phase separation in epitaxial layers as described by spinodal decomposition.^{1–8} Phase separation in III–V ternary alloys can occur when local variations in cation concentrations and strain favor “up-hill” diffusion as a means of lowering the free energy of the alloy. The growth of short period strained superlattices likewise creates more interfaces, which enhance the number of cations diffusing laterally and further facilitates phase separation. This strain induced lateral ordering (SILO) method has been utilized to form quantum wire (QWR) arrays in bilayer superlattices (BSLs) of $(\text{InP})_2/(\text{GaP})_2$ and $(\text{GaAs})_2/(\text{InAs})_2$ grown on a GaAs(001) or InP(001) substrate, respectively.^{9–14} The optical properties of the QWR structures have been studied with photoluminescence (PL),^{9,10} photoreflectance (PR),¹¹ and cathodoluminescence (CL)^{12–14} techniques. A large polarization anisotropy was found to be consistent with the existence of QWRs.^{9,13} It is of paramount importance to study the optical properties of SILO QWRs in order to understand the interplay between the microstructure, local optical properties, and growth conditions. These QWRs further exhibit good potential for applications in optoelectronics such as for light-emitting diodes, high-speed optical switches, and low-threshold, high-gain QWR lasers.^{10,15}

In a previous study, we used CL techniques to examine the nonlinear optical properties of $(\text{InP})_2/(\text{GaP})_2$ BSLs on GaAs(001) substrates grown by metal–organic chemical vapor deposition (MOCVD).^{12–14} The combination of linearly

polarized CL spectroscopy, CL imaging and time-resolved CL enabled a study of the nonlinear optical properties and carrier relaxation kinetics in the QWR structure.^{12–14} In this study, using these CL techniques, we examine the optical and structural properties in the QWR structures grown by gas source molecular beam epitaxy (MBE). Based on the observed optical and structural properties, we show that MBE enables a good control of the growth of short period superlattices used to produce QWR arrays. The spatial homogeneity of the QWR structures was studied with CL imaging and a depth dependent CL study. Misfit dislocations and spot defects were observed in the CL imaging. The QWR luminescence exhibits multiple components; their origin and carrier relaxation into regions giving rise to these components are studied with time-resolved CL.

II. EXPERIMENTAL DETAILS

A schematic diagram of the sample structure is shown in Fig. 1. BSLs of $(\text{InP})_2/(\text{GaP})_2$ were grown on a 3000 Å thick $\text{In}_{0.49}\text{Ga}_{0.51}\text{P}$ buffer layer, which is lattice matched to the underlying GaAs(001) substrate. Details of the gas source MBE system and the BSL growth procedures have been previously reported.^{6,9,10} Two samples were grown for this study and labeled as 2067 and 2339, respectively. Each sample consists of four $(\text{InP})_2/(\text{GaP})_2$ BSL regions, each separated by 180-Å-thick $\text{In}_{0.49}\text{Ga}_{0.51}\text{P}$ barriers and is capped with a 300-Å-thick AlInP layer. Each of the four $(\text{InP})_2/(\text{GaP})_2$ BSL regions are ~147 and ~102 Å thick for samples 2067 and 2339, respectively. The rate of the BSL growths was ~1 ML/s. The lateral composition modulation in the bilayer structure occurred during the gas source MBE growth, and the modulation direction is along the [110] direction. The composition of In varies from ~0.33 in the Ga-rich region to ~0.67 in the In-rich region, with periods of

^{a)} Author to whom correspondence should be addressed; Electronic mail: danrich@almaak.usc.edu

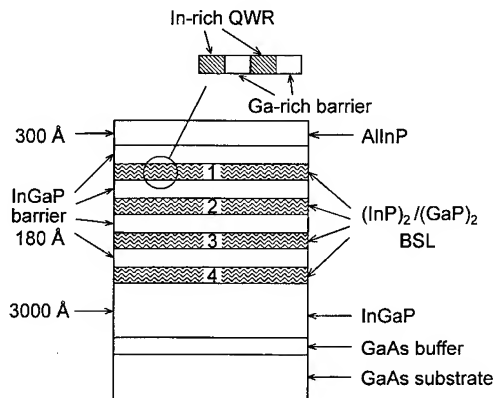


FIG. 1. Schematic diagram of $(\text{InP})_2/(\text{GaP})_2$ bilayer superlattice structure.

~ 125 and ~ 100 Å for sample 2067 and 2339, respectively.⁹ Therefore, quantum wire arrays were formed with cross-sectional areas of ~ 147 Å \times 63 Å and ~ 102 Å \times 50 Å with a two-dimensional (2D) quantum confinement in the growth and lateral ordering directions.

The CL experiments were performed with a modified JEOL-840A scanning electron microscope with polarization detection capability described previously.¹²⁻¹⁴ In polarization measurements, light with electric-field E parallel to the $[110]$ or $[1\bar{1}0]$ was detected. The luminescence signal from the sample, whose temperature was maintained at ~ 87 K by a liquid-nitrogen cryogenic specimen stage, was dispersed by a $1/4$ m monochromator and detected by a cooled GaAs:Cs photomultiplier tube with a spectral resolution of ~ 1 nm. In time-resolved experiments, the method of delayed coincidence in an inverted single photon counting mode was used.¹⁶ Electron-beam pulses of 50 ns width with a 1 MHz repetition rate were used to excite the samples. This enabled the measurement of time-delayed CL spectra with a temporal resolution of ~ 100 ps. The e-beam energy, E_b , was varied systematically from 3 to 35 keV in the experiments to obtain a tunable electron range and probe depth.

III. RESULTS AND DISCUSSION

A. Excitation-dependent CL spectra

CL spectra of both samples are shown in Fig. 2. The broad peaks located in the $680 \leq \lambda \leq 750$ nm range are due to emission from the QWR array.⁹⁻¹⁴ The CL spectral line shapes are observed to vary with beam energy, particularly for sample 2067, which shows a marked enhancement in the high-energy side of the line shape for higher beam energies. Spectra taken at $E_b = 3$ keV, for both samples, show a line shape very similar to that of PL obtained with an argon-ion laser (not shown here). The polarization anisotropy of the integrated CL intensity over the $680 \leq \lambda \leq 750$ nm range were measured, and $I_{\parallel}/I_{\perp} \approx 3$, where I_{\parallel} and I_{\perp} are the integrated CL intensities taken with the polarizer orientated so that the electric-field E of the detected light is parallel and perpendicular to the quantum wire, respectively. This polarization ratio is about twice as large as compared to samples grown by MOCVD ($I_{\parallel}/I_{\perp} \approx 1.5$).^{12,13} This may be due to a higher

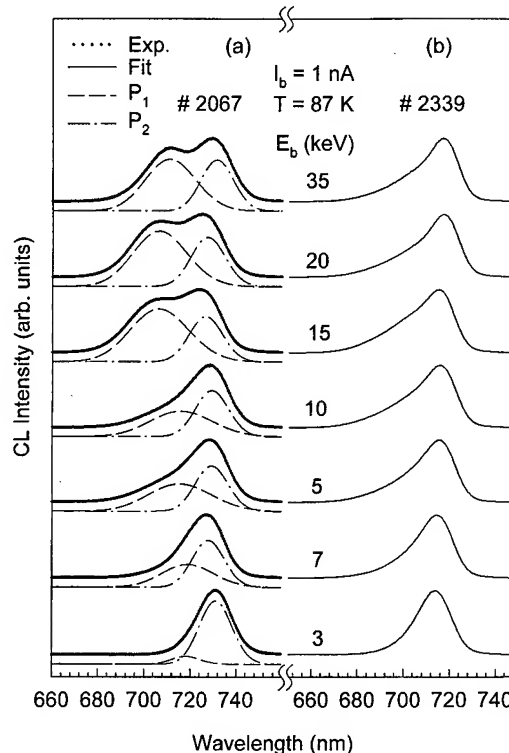


FIG. 2. Stack plots of CL spectra taken with E_b varying from 3 to 35 keV for samples 2067 (a) and 2339 (b). In (a), spectra are fitted with a sum of two Gaussian functions. Solid lines show the fitting results with the decomposed P_1 (dotted lines) and P_2 (dash lines) components offset below each spectrum.

degree of ordering and narrower QWR thickness in the MBE-grown structures, whose compositional modulation period (~ 100 Å) is much shorter than that observed for the MOCVD-grown structures (~ 800 Å).¹²⁻¹⁴ The substrate temperature for the MOCVD¹² and MBE⁶ growths were maintained at ~ 650 and ~ 500 °C, respectively. Owing to the temperature difference between the MBE and MOCVD growth, we should, therefore, expect differences accordingly in the composition and spatial wavelength of the resultant QWR regions.⁶

The e-beam probing depth dependence of the CL spectra was studied and a stack plot of CL spectra taken under electron-beam energies E_b ranging from 3 to 35 keV is shown in Fig. 2. We found that in sample 2067 a higher energy luminescence feature emerged for $E_b \geq 10$ keV, and its intensity increased as the electron penetration depth increased. A smaller increase in the CL intensity from the high-energy side of the line shape is observed for sample 2339 in Fig. 2 as the E_b increases. A probe current dependence experiment showed that for a fixed $E_b = 5$ keV, no significant change in the CL spectra line shape occurred in both samples as the probe current increases from 50 pA to 5 nA. Thus, the appearance of these high-energy peaks cannot be attributed directly to an enhanced excitation density and band filling for higher e-beam energies. The two luminescence features evidently arise from separate regions at different depths of the sample. Likewise, in a previous CL

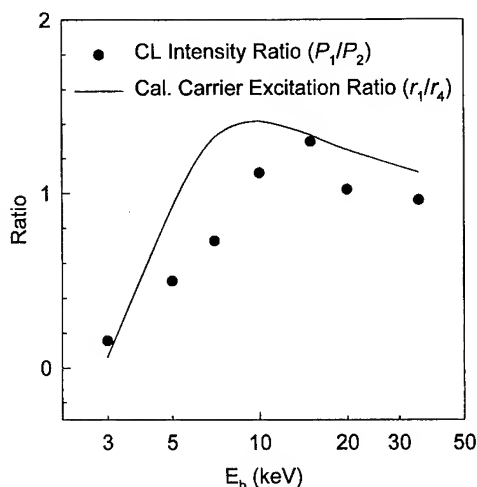


FIG. 3. Intensity ratio of P_1 and P_2 components (filled symbols) and ratio of the calculated carrier excitation rate, r_1/r_4 , in the first and fourth BSL layers (solid line) as a function of beam energy.

study, we observed the presence of two or more components in the QWR emission and suggested that they arise from two different QWR widths and/or In compositions.¹⁴ To further quantify the behavior of the CL spectra, we have decomposed the CL line shapes of sample 2067 into the sum of two Gaussian peaks, P_1 and P_2 . The fitting results (solid line) are shown in Fig. 2(a) with the two components offset below. The intensity ratio of P_1 and P_2 is shown as a function of beam energy in Fig. 3. For comparison, we show the ratio of the calculated $e-h$ excitation rates, r_i , between QWR regions 1 and 4 (i.e., r_1/r_4 , solid line in Fig. 3). We have used a simple one-dimensional depth-dose $e-h$ pair creation function, ignoring carrier diffusion, in this calculation.¹⁷ While a lateral diffusion length of $\sim 0.5 \mu\text{m}$ is typical for the InGaP material system,¹⁸ the effective diffusion length along the growth direction will be reduced as a result of the BSL regions impeding vertical carrier transport. Similar effects have been observed in the related InGaAs/GaAs multiple quantum well system.¹⁹

It is, thus, reasonable to conclude that P_1 and P_2 originate from the inner (deeper) and outer (shallower) QWR regions, respectively. It is likely that variations in the lateral QWR width and/or In composition modulation occurred during the SILO process. Previous cross-sectional transmission electron microscopy results have shown that the vertical and lateral dimensions of the QWR depend on the location in the BSL region and total film thickness.^{6,9} The inner BSL layers have a narrower lateral QWR than that of the outer BSL layers, owing to a strain accumulation effect during growth, which results in a layer-dependent bulk diffusion and phase separation.^{6,9} The presence of misfit dislocations (see Sec. III C) indicates the buildup of a net global strain and subsequent strain relaxation during the latter stages of the growth.

Using a transfer matrix method to calculate the ground state excitonic transition energies in the QWRs,¹³ we estimate that the P_1 and P_2 energy difference of 47 meV results

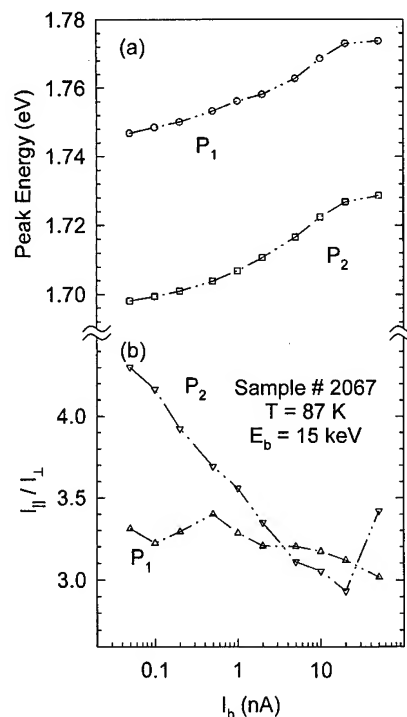


FIG. 4. Peak energy E_p (a) and polarization ratio I_{\parallel}/I_{\perp} (b) as a function of probe current for luminescence components P_1 and P_2 for sample 2067.

in a maximum In composition variation of $\sim 6\%$ for a fixed QWR size. For a fixed In composition in the QWRs, no reasonable variation in the QWR width or height would account for a 47 meV variation in the ground state excitonic transition energy. Therefore, we conclude that the energy splitting between P_1 and P_2 and QWR width fluctuations are also accompanied by differences in the In composition between the inner and outer BSL layers.

B. Band filling and nonlinear optical effects in QWRs

Band filling effects in these QWR structures were also examined. In sample 2067, we observed blueshifts in the peak positions of P_1 and P_2 as the excitation density increased. By increasing the probe current, I_b , from 50 pA to 10 nA, components P_1 and P_2 showed a nearly identical shift of ~ 30 meV, as shown in Fig. 4. However, the polarization ratio I_{\parallel}/I_{\perp} varies differently for P_1 and P_2 as also shown in Fig. 4. The ratio I_{\parallel}/I_{\perp} reduces for P_2 as the probe current increases, but remains nearly constant for P_1 . The change of polarization anisotropy as a result of band filling has been studied previously with a $\mathbf{k} \cdot \mathbf{p}$ model that takes both 2D quantum confinement and the coherent strain into account, showing important nonlinear optical effects in these samples.¹³ Applying this model to these sample structures, we find that the polarization ratio should likewise reduce as the excitation density increases for these samples. The reason for the slower change in the polarization ratio of P_1 is, however, unclear, and may relate to differences in heavy- and light-hole characters between different QWR regions.¹³

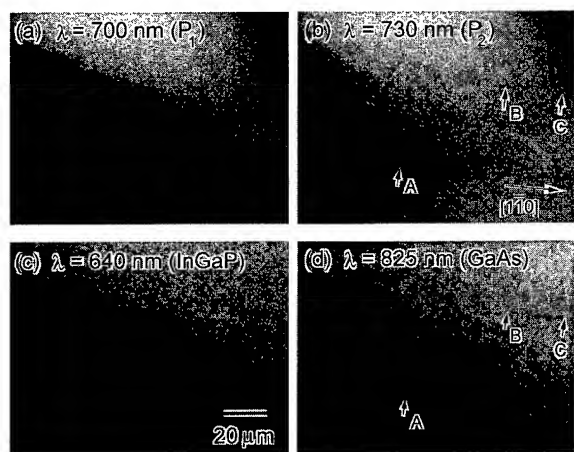


FIG. 5. Monochromatic CL images taken from sample 2067 with wavelength corresponding to BSL luminescence components P_1 (a) and P_2 (b), $\text{In}_{0.49}\text{Ga}_{0.51}\text{P}$ bulk emission (c) and GaAs substrate emission (d).

C. Monochromatic CL imaging of point defects and DLDs

To further investigate the complex origin of the two components in QWR emission, CL images were taken with wavelengths set to the peak wavelength positions for P_1 and P_2 , $\text{In}_{0.49}\text{Ga}_{0.51}\text{P}$ bulk, and GaAs substrate emissions, as shown in Fig. 5. An electron beam with energy $E_b = 15$ keV was used to obtain these CL images for sample 2067. A spotty pattern is observed with spot positions correlating in all four images. This is similar to that observed in MOCVD-grown structures.¹² The spatial correlation among these images indicates that these defects arise from the GaAs substrate or were formed during the initial stages of growth of $\text{In}_{0.49}\text{Ga}_{0.51}\text{P}$ on the GaAs substrate. These defects may form as a result of the agglomeration of In on the GaAs surface, as previously observed during InGaP/GaAs growth.¹² In addition, dark line defects (DLDs) were observed in the images corresponding to the P_2 and GaAs luminescence in Fig. 5. Such DLDs remarkably were not detected in the images for P_1 and $\text{In}_{0.49}\text{Ga}_{0.51}\text{P}$ emission. As discussed above, P_2 dominates the luminescence when the excitation is close to the surface and QWR region 1. The absence of DLDs in the intermediate regions denoted by P_1 and InGaP luminescence (i.e., the 700 and 640 nm images, respectively) is at first glance troubling. However, a close inspection of the 730 and 825 nm images reveals that the DLDs are not directly spatially correlated. Rather, some DLDs appear as finite parallel segments whose positions in the 730 and 825 nm images are displaced along the $[1\bar{1}0]$ direction. Consider, for example, positions marked by arrows A, B, and C in these images. Arrows A and C point to DLD segments in the 825 nm image, which are absent in the 730 nm image. The converse is true for the region marked by arrow B, denoting a DLD segment in the 730 nm image. Such a behavior reflects a stepping sequence in the nucleation and propagation of 60° misfit dislocations, and has been observed in InGaAs/GaAs MQWs.^{20,21}

In the standard approach to dislocation formation in

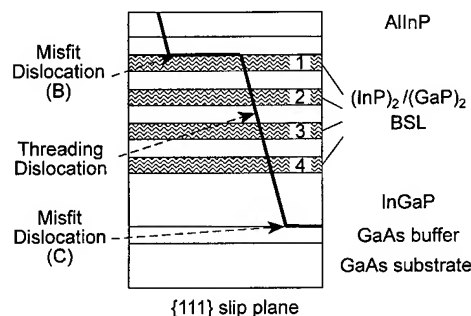


FIG. 6. Schematic showing stepping sequence of 60° dislocations in the BSL structure. In this model, the threading dislocation segments elongate (bend over) into misfit dislocations at the GaAs/InGaP (bottom) and BSL/InGaP (top) interfaces. The diagram shows a $\{111\}$ slip plane. Misfit dislocation segments B and C correspond to those labeled in Fig. 5.

strained III-V systems,²² the substrate is considered to have a fixed number of threading dislocation sources and substrate surface defects, which act as heterogeneous nucleation sites for 60° threading dislocations. Once the thickness of the InGaP and $(\text{InP})_2/(\text{GaP})_2$ BSL layers has reached a value where the strain force is equal to the dislocation line tension, the dislocation will glide laterally on a $\{111\}$ slip plane, resulting in the elongation of the dislocation in an interfacial plane. The length of the misfit segment during growth is controlled by several factors, including frictional forces, impurities, defects, and surface steps, which can effect the eventual pinning of the threading dislocation. A pinned threading dislocation will replicate during growth. Once sufficient strain is attained in the subsequent growth, the dislocation can again elongate on a different layer of the BSL. The current CL data suggests a simple stepping sequence schematically depicted in Fig. 6. Misfit dislocations form at the lower GaAs/InGaP interface, after which the pinned threading segments again elongate into a second misfit dislocation near the top BSL 1 interface. This model explains the segmentation and correlation of the DLD contrast in the 730 and 825 nm images of Fig. 5. To our knowledge, this is the first observation of a dislocation stepping behavior in the InGaP/GaAs system.

In sample 2339, the CL images (Fig. 7) showed the characteristic spotty pattern, but no evidence of DLD formation was observed. Also, the variation in line shape with beam energy was substantially reduced compared to the behavior of sample 2067 above. These results strongly suggest that the strain in 2339 is below that necessary to induce the formation of misfit dislocation since the thickness of the BSL in sample 2339 is smaller than that in sample 2067. Therefore, in SILO QWR structures, we demonstrate an apparent relationship between exceeding the critical thickness for misfit dislocation and DLD formation and the presence of multiple components in the QWR luminescence.

D. Time-delay CL measurements of the carrier relaxation, reemission, and diffusion

In order to examine the relaxation and collection of carriers into the QWRs, we have performed a time-resolved CL

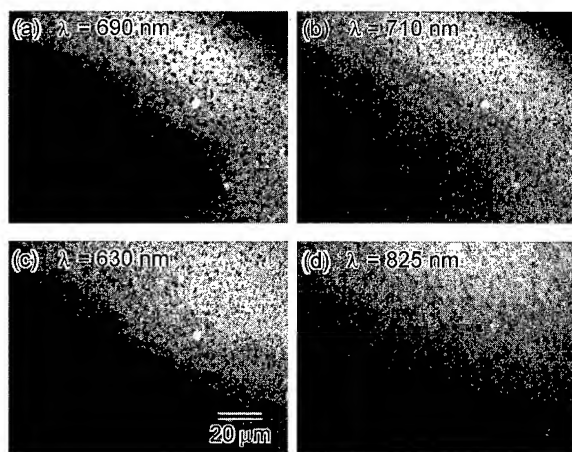


FIG. 7. Monochromatic CL images taken from sample 2339 with wave-length $\lambda=690$, 710, 630, and 825 nm, respectively.

study at $T=87$ K. Shown in Fig. 8 is a stack plot of CL spectra taken from sample 2067 for various time windows with $E_b=15$ keV and $I_b=10$ nA. During the onset phase (windows O1–O5), a broad feature for $680 \leq \lambda \leq 750$ nm is observed with no distinct peaks, indicating a relatively uniform capture into all four QWR regions. A small narrowing and redshifting of the spectra is observed during the progression from the O1 to O6 time windows. The peak labeled *in pulse* represents a measurement in the center of the 50 ns excitation pulse during which a near-steady-state $e-h$ generation and recombination rate has been reached. A rapid

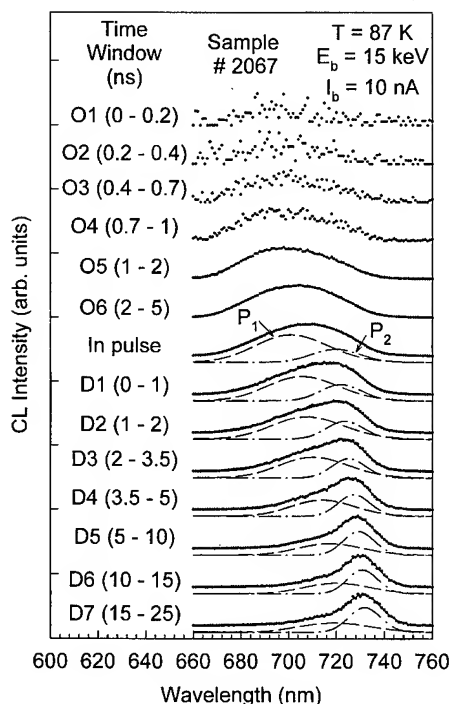


FIG. 8. Stack plot of time-delay CL spectra various onset (Oi) and decay (Di) time windows (for sample 2067). All spectra are renormalized to have about the same maximum peak height.

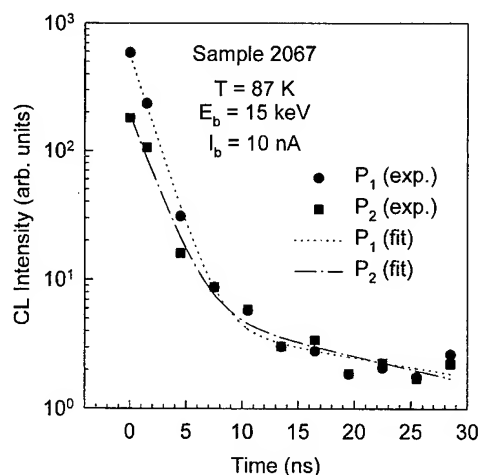


FIG. 9. Intensities of BSL emission components P_1 and P_2 as a function of time. At $t=0$, the electron beam is turned off. Biexponential fits are shown with dotted and dash lines for P_1 and P_2 , respectively.

narrowing and redshifting is observed during the decay phase, D1–D7 windows in Fig. 8. The CL intensity of the two CL components P_1 and P_2 versus time is obtained by decomposing the time-delayed CL spectra for each time window shown in Fig. 9. The CL intensity decay versus time is shown in Fig. 9, where dot and dash lines represent biexponential fits to the data.²³ The initial decay time constant is 1.6 ns for P_1 and 1.8 ns for P_2 . The faster decay rate for P_1 and the reversal of P_1 and P_2 intensities indicate that channels for the diffusive transport of carriers between the P_1 and P_2 levels (QWR regions 4 and 1) exist. This should not be surprising, given that the distance between regions 1 and 4 is ~ 830 Å, substantially less than the diffusion length of ~ 0.5 μm . Thermal reemission from the QWR regions before recombination will occur, leading to an eventual carrier transfer into the lowest energy (P_2 -like) states for transients exhibiting the longest decay times. The redshift of both P_1 and P_2 during the decay phase is ~ 30 meV (determined from fits not shown) and this is due to the reduction of quasi-Fermi-level difference as excess carrier density decreases with time. The energy shifts in the time-delayed CL spectra are complimentary to those observed in the excitation-dependent results of Fig. 4(a).

IV. SUMMARY AND CONCLUSIONS

In conclusion, we have examined the nonlinear optical properties and structural properties using the cathodoluminescence techniques. In comparison to the MOCVD-grown structure previously studied,^{12–14} the gas source MBE-grown structures in this study show a higher degree of lateral ordering and narrower QWRs. The narrower QWRs in MBE suggest that surface reaction kinetics and the mass transfer rates play an important role in the lateral diffusion and phase separation of (InP)₂/(GaP)₂. Two distinct components, P_1 and P_2 , were observed in the sample with the thicker QWR regions. The beam energy dependence of the CL spectral line shape indicated variations in the In composition modulation

and QWR lateral dimensions between the inner and outer (InP)₂/(GaP)₂ BSL regions. The outer BSL has a ~6% higher In composition modulation as a result of a strain accumulation effect.^{6,9} Strong phase-space filling and nonlinear optical effects were observed for these samples, consistent with the narrowing of the density of states in QWRs.¹³ Likewise, a strong excitation dependence of the polarization anisotropy was observed, also consistent with a QWR-like behavior.¹³ CL imaging, however, revealed defects in a spotty pattern that are tied to the initial stages of InGaP growth on GaAs. Also, misfit dislocations were found to form for the sample with a thicker BSL region. Monochromatic CL imaging revealed that misfit dislocations and DLDs formed in a stepping sequence with finite segments close to the GaAs/InGaP interface and uppermost BSL layers (BSL region 1). Time-delayed CL spectra were acquired to examine the carrier capture and relaxation dynamics. A uniform capture of carriers into the four QWR regions was observed during the onset time windows. A rapid transfer of carriers between the QWR regions occurs on a sub-ns time scale. This was observed during the decay phase, as carrier reemission, diffusion, and thermalization occurs between inner and outer BSL regions, i.e., over distances of ~0.1 μm. The quantum efficiency of QWR lasers depends on the carrier collection efficiency of the QWRs. Future time-resolved CL studies of these SILO QWRs will focus on the temperature dependence of the carrier relaxation, band filling, and polarization anisotropy in order to optimize the carrier collection in the QWR at room temperature.

ACKNOWLEDGMENTS

This work was supported by the U.S. Army Research Office (Grant Nos. DAAH04-94-G-0260 and DAAH04-95-1-0386), the National Science Foundation (Grant Nos. ECS-94-09122, ECD-89-43166, and DMR-89-2-538), the Joint Services Electronics Program (N00014-96-1-0129), and the Office of Naval Research University Research Initiative Program (N00014-92-J-1519KC).

- ¹J. W. Cahn, *Acta Metall.* **9**, 975 (1961).
- ²G. B. Stringfellow, *J. Cryst. Growth* **27**, 21 (1974).
- ³P. Henoc, A. Izrael, M. Quillec, and H. Launois, *Appl. Phys. Lett.* **40**, 963 (1982).
- ⁴F. Glas, *J. Appl. Phys.* **62**, 3201 (1987).
- ⁵A. A. Mbaye, L. G. Ferreira, and A. Zunger, *Phys. Rev. Lett.* **58**, 49 (1987).
- ⁶K. C. Hsieh and K. Y. Cheng, *Mater. Res. Soc. Symp. Proc.* **379**, 145 (1995).
- ⁷A detailed discussion of the relevance of Refs. 1–5 towards understanding strain-induced lateral ordering in III–V alloys can be found in Ref. 6.
- ⁸K. C. Hsieh, K. Y. Cheng, Y. L. Hwang, T. Zhang, and R. M. Kolbas, *Appl. Phys. Lett.* **68**, 1790 (1996).
- ⁹S. T. Chou, K. Y. Cheng, L. J. Chou, and K. C. Hsieh, *J. Appl. Phys.* **78**, 6270 (1995); A. C. Chen, A. M. Moy, P. J. Pearah, K. C. Hsieh, and K. Y. Cheng, *Appl. Phys. Lett.* **62**, 1359 (1993).
- ¹⁰P. J. Pearah, A. C. Chen, A. M. Moy, K. C. Hsieh, and K. Y. Cheng, *IEEE J. Quantum Electron.* **30**, 608 (1994).
- ¹¹A. Mascarenhas, R. G. Alonso, G. S. Horner, S. Froyen, K. C. Hsieh, and K. Y. Cheng, *Phys. Rev. B* **48**, 4907 (1993).
- ¹²Y. Tang, K. Rammohan, H. T. Lin, D. H. Rich, P. Colter, and S. M. Vernon, *Mater. Res. Soc. Symp. Proc.* **379**, 165 (1995).
- ¹³Y. Tang, H. T. Lin, D. H. Rich, P. Colter, and S. M. Vernon, *Phys. Rev. B* **53**, R10 501 (1996); D. H. Rich, Y. Tang, and H. T. Lin, *J. Appl. Phys.* **81**, 6837 (1997).
- ¹⁴D. H. Rich and Y. Tang, *Appl. Phys. Lett.* **69**, 3716 (1996).
- ¹⁵E. M. Stellini, K. Y. Cheng, P. J. Pearah, A. C. Chen, A. M. Moy, and K. C. Hsieh, *Appl. Phys. Lett.* **62**, 458 (1993).
- ¹⁶D. Bimberg, H. Munzel, A. Steckenborn, and J. Christen, *Phys. Rev. B* **31**, 7788 (1985); H. T. Lin, D. H. Rich, A. Konkar, P. Chen, and A. Madhukar, *J. Appl. Phys.* **81**, 3186 (1997).
- ¹⁷We integrate the depth-dose curve $g(z)$ over the first and fourth layers to estimate the ratio of the $e-h$ generation rates, as described in T. E. Everhart and P. H. Hoff, *J. Appl. Phys.* **42**, 5837 (1971).
- ¹⁸M. Prasad, O. E. Martinez, C. S. Menoni, J. J. Rocca, J. L. A. Chilla, M. J. Hafich, and G. Y. Robinson, *J. Electron. Mater.* **23**, 359 (1994).
- ¹⁹K. Rammohan, H. T. Lin, D. H. Rich, and A. Larsson, *J. Appl. Phys.* **78**, 6687 (1995).
- ²⁰J. Wang, J. W. Steeds, and D. A. Woolf, *Philos. Mag. A* **65**, 829 (1992).
- ²¹D. H. Rich, H. T. Lin, Y. Tang, K. Rammohan, and A. Larsson, *Scanning Microsc. Suppl.* (in press).
- ²²J. W. Mathews, A. E. Blakeslee, and S. Mader, *Thin Solid Films* **33**, 253 (1976).
- ²³P. Dua, S. L. Cooper, A. C. Chen, and K. Y. Cheng, *Appl. Phys. Lett.* **69**, 2261 (1996).

Optical properties of semiconductor nanostructures

T. L. Reinecke^{a)}

Naval Research Laboratory, Washington, DC 20375

P. A. Knipp

Department of Physics and Computer Science, Christopher Newport University, Newport News, Virginia 23606

S. N. Walck

Naval Research Laboratory, Washington, DC 20375

(Received 5 March 1997; accepted 29 March 1997)

Boundary element methods for calculating the single-particle electron and photon states of semiconductor quantum wires and quantum dots are discussed, and variational treatments for calculating exciton properties are given. Calculations of single-particle carrier states, of exciton binding energies and of confined photon states are compared with experimental data obtained on lithographically formed semiconductor wires and dots. © 1997 American Vacuum Society. [S0734-211X(97)12304-6]

Low-dimensional semiconductor systems, particularly quantum wires and quantum dots, are attracting considerable attention recently, in part, because they exhibit novel physical properties and also because of potential applications involving them. In recent years, a number of innovative techniques have been developed to grow or to fabricate and to study experimentally a variety of quantum wire and quantum dot structures having different geometries and potentials. The optical properties of these systems are especially useful in giving detailed information about their microscopic physics.

To date, the basic physical properties of these systems are less understood than are those of bulk and quantum well systems. Except for certain highly symmetrical cases, such as spherical quantum dots and cylindrical wires, the geometries of realistic structures generally result in the equations of motion and boundary conditions governing their elementary excitations, carriers, phonons, and photons, being nonseparable. These systems are often discussed using simplified models, but it often is difficult to assess the validity of such approaches. Alternately, numerical approaches can be used. The most useful methods, to date, for studying these nanostructures are finite difference calculations¹ and boundary element methods.² In finite difference calculations, the equations of motion are straightforwardly discretized on a mesh of the nonseparable coordinates and are solved by matrix techniques. This method is particularly useful where electronic self-consistency must be considered and alternative methods are not available. In the following, we give a summary of the boundary element method, and some comparisons of calculations and experimental data for single-particle electron and hole states, for excitons, and for confined photons in quantum wires and quantum dots.

In recent years, we have developed a set of boundary element methods for calculating the phonon,³ electron,⁴ and photon⁵ properties of low-dimensional structures. These

methods are based on ideas from traditional scattering theory and have been found to be especially efficient in treating systems that are composed of several regions each having uniform properties. In this method, Green's function techniques are used to transform, for example, a problem involving a quantum wire having two nonseparable variables running over its cross-sectional area into a one-dimensional problem involving coordinates only around the circumference of the wire. The resulting numerical calculations then generally are considerably more efficient than are those from finite differences methods.

In order to illustrate the boundary element method, we consider bound electronic states described by Schrödinger's equation, which is written in the effective mass approximation

$$\left[\frac{-\hbar^2}{2} \nabla \cdot \frac{1}{m(\mathbf{r})} \nabla + V(\mathbf{r}) - E \right] \psi(\mathbf{r}) = 0, \quad (1)$$

where $V(\mathbf{r})$ and $m(\mathbf{r})$ are the potentials and the masses for several regions, each of which is taken to be uniform. The corresponding Green's function is determined by

$$\left[\nabla^2 + \frac{2m}{\hbar^2} (E - V) \right] G(\mathbf{r}, \mathbf{r}'; E) = \delta(\mathbf{r} - \mathbf{r}'). \quad (2)$$

For regions having uniform properties, the Green's functions are well known and can be obtained relatively simply. From Eqs. (1) and (2) and the boundary conditions on the wave functions and their normal derivatives at the boundaries, the two central equations are obtained. First, the wave function in all space is given in terms of values on the boundaries by the integral

$$\psi(\mathbf{r}') = \int_S dA [-\phi(A) G(A, \mathbf{r}'; E) + \psi(A) \partial_A G(\mathbf{r}, \mathbf{r}'; E)], \quad (3)$$

and second, the wave function and its derivative on the boundary are obtained from the integral equation

^{a)}Electronic mail: tom.reinecke@nrl.navy.mil

TABLE I. Convergence of the two lowest lying levels of a stadium structure as a function of the number of boundary elements N . The stadium shape is given in Fig. 1. Here, it is 25 nm×50 nm in size with a band offset of 10 meV, and the electron mass is taken to be 0.0665.

N	E_0	E_1
4	4.8823	8.7019
8	4.8095	8.7019
16	4.8042	8.6403
32	4.8028	8.6334
64	4.8024	8.6312
128	4.8023	8.6308

$$\int_S dA [B(\mathbf{A}, \mathbf{A}') \phi(\mathbf{A}') + C(\mathbf{A}, \mathbf{A}') \psi(\mathbf{A}')] = 0. \quad (4)$$

Here, S is the surface bounding the region V , \mathbf{A} indicates a point \mathbf{r}_A on S , $\phi(\mathbf{A}) = \partial_A \psi(\mathbf{r})$, $\partial_A = \hat{n}_A \cdot \partial / \partial \mathbf{r}$, and \hat{n}_A is the normal pointing out of V at \mathbf{r}_A . The integral in Eq. (4) then is discretized along the boundary S , and the equation is solved numerically. Spectral densities and scattering cross sections, which are the quantities often of direct experimental interest, are given by equations analogous to Eqs. (1) and (2) above, and therefore, they can be calculated directly in this method rather than by summing over individual wave functions.

In Table I we give results for the convergence of the energies of the two lowest lying electronic states of a stadium structure with a potential offset of 10 meV. This structure is composed of two parallel straight lines and closed by two semicircles, as indicated in Fig. 1, and for it, Schrödinger's equation is nonseparable. Such structures are of interest in connection with the study of scarred wave functions and the onset of quantum chaos. For comparison, a higher lying electronic state with energy of 184 meV in a stadium with a potential offset of 190 meV is shown in Fig. 1. Approximately 50 boundary elements were needed for its calculation.

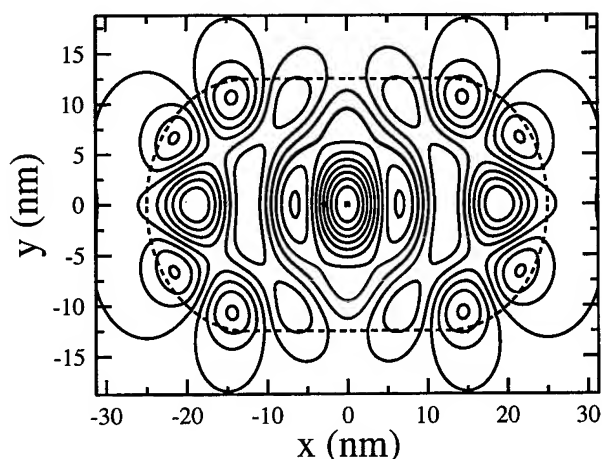


FIG. 1. Contour plot of the confined electron state with energy 184 meV of the stadium structure 25 nm×50 nm in size with a potential offset of 190 meV. The dashed line gives the shape of the stadium.

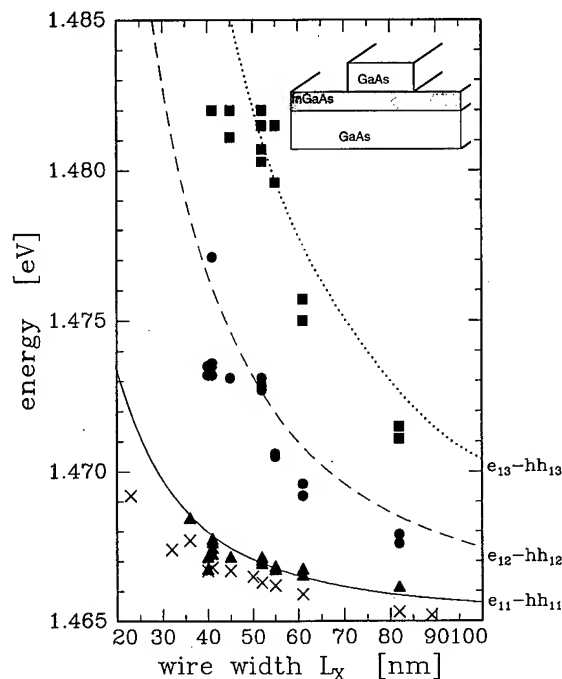


FIG. 2. Photoluminescence excitation data for and calculations of (solid lines) electron-hole transition energies vs wire width for modulated barrier GaAs/In_{0.10}Ga_{0.90}As quantum wires with a quantum well of 5 nm width. The inset shows a sketch of the modulated barrier structure.

culcation. The convergence of the energies in these calculations with the number of boundary elements is quite good.

Calculations for realistic models can be used to understand experimental results for the systems. An interesting case involves barrier modulated structures,⁶ which are sketched in the inset in Fig. 2. For them, a barrier layer of GaAs covering an InGaAs quantum well is selectively etched off, and for sufficiently narrow sizes a quasi-one-dimensional quantum wire for electrons and holes is formed beneath the remaining barrier in the quantum well layer. These structures have high optical quality because the etching does not extend into the quantum well. From calculations of the electron and hole states in these structures, we find for decreasing wire widths that the electron and hole wave functions are increasingly pushed into the higher gap GaAs substrate, thus, increasing their energies. This behavior is a result of the nonseparable character of the wave function for this case, and it must be taken into account in order to understand quantitatively the size dependence of the lateral quantization in them. In Fig. 2, photoluminescence excitation results⁷ for electron-hole recombination as a function of wire width are shown along with results of these calculations. In these structures, three electron and three hole states are confined laterally. Here, the wire widths are those obtained directly from scanning electron microscopy (SEM) measurements. The calculations are seen to be in good agreement with the data with no adjustments for damage, such as optical dead layers, thus, indicating that the structures are of good optical quality.

Results from magnetoluminescence investigations of

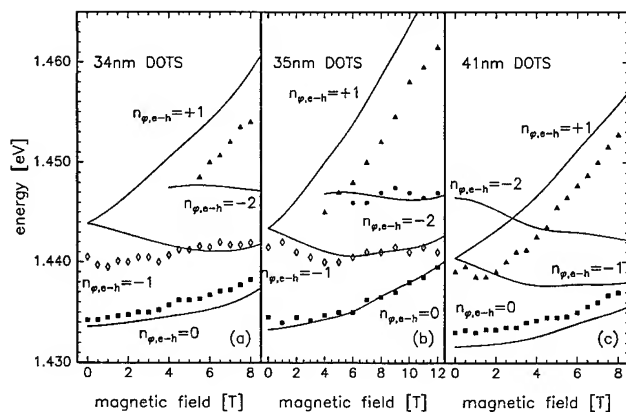


FIG. 3. Magnetoluminescence data for and calculations of (solid lines) electron-hole recombination energies in modulated barrier GaAs/In_{0.10}Ga_{0.90}As quantum dots with a quantum well of 5 nm width.

electron-hole recombination energies of modulated barrier quantum dots⁸ are shown in Fig. 3 with results from calculations for these structures. SEM images for these dots indicate that they are, to a good approximation, cylindrical in shape, and therefore, the angular momentum around the cylinder axis is a good quantum number. At zero magnetic field, the ground electron and hole states are orbitally nondegenerate, and the first excited states are twofold degenerate, corresponding to two orientations of the angular momentum. For nonzero magnetic fields, the degeneracy of the first excited state is lifted, and the states are split. At very high magnetic fields, the states go over to the Landau levels of a quantum well. These calculations are in good qualitative and quantitative agreement with the data, and these results provide the most detailed picture to date of the ground and excited states of a quantum dot system.

The effects of confinement on exciton binding energies has been one of the most active topics involving low-dimensional systems, and it is of interest for applications. This behavior has been observed in quantum wells, but to date, there has been no systematic investigation of it for lower-dimensional systems. This is in part because for them, the exciton binding energy cannot be obtained using low-field extrapolations of magneto-optical results. We have made detailed studies of exciton binding energies for modulated barrier and deep etched quantum wires and quantum dots and have compared them with corresponding experimental results for In_{0.10}Ga_{0.90}As/GaAs quantum dots and quantum wires.⁹ These structures provide results for systems with widely varying sizes and potential heights. The experimental results for the binding energies were obtained using a method¹⁰ in which the excitonic diamagnetic shifts are measured, and the binding energies are obtained from them using a relation between the diamagnetic shift and the binding energy for excitons in bulk systems with anisotropic bulk masses. These results for modulated barrier GaAs/InGaAs wires and dots are shown in Fig. 4.

In these structures, the relative and center-of-mass coordinates of the excitons in the plane perpendicular to the mag-

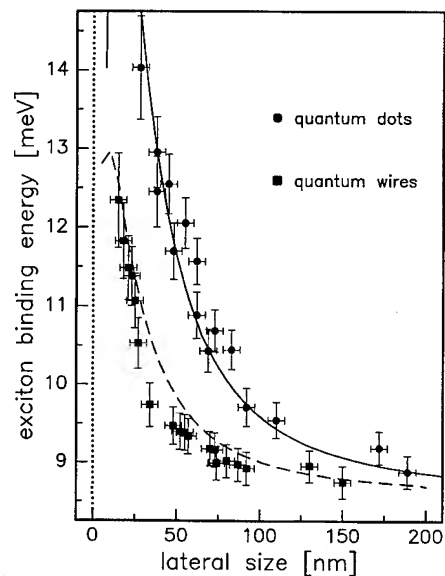


FIG. 4. Exciton binding energies from magnetoluminescence studies on modulated barrier quantum wires and quantum dots as functions of size. Solid curves give calculations described in the text.

netic field in these experiments do not separate. For these systems, we perform a canonical transformation on the Hamiltonian for the interacting electron and hole of the form

$$H' = e^{iG} H e^{-iG}, \quad (5)$$

where

$$G = -\frac{eB}{2\hbar c} (\lambda_1 x_e y_h - \lambda_2 x_h y_e - \lambda_3 x_e y_e + \lambda_4 x_h y_h). \quad (6)$$

Here λ_i , $i=1,4$, are variational parameters, which are chosen to minimize the exciton energy for each structure at each magnetic field. The exciton binding energies are calculated variationally using the two parameter functions

$$\Psi_{\text{ex}} = \zeta_e(\cdots) \zeta_h(\cdots) \exp\{-[\alpha(x_e - x_h)^2 + y_e - y_h^2] + \beta(z_e - z_h)^2\}^{1/2}. \quad (7)$$

Here, $\zeta_e(\cdots)$ and $\zeta_h(\cdots)$ are the envelop functions for the electron and hole subbands. For these nonseparable structures, the subband functions are calculated numerically as discussed above, and the electron-hole correlation in the exciton is calculated variationally by minimizing the energy with respect to α and β . Variational functions of the form in Eq. (7) have been found to give good results for exciton binding energies in quantum wells.

The diamagnetic shifts that we calculate are in good agreement with the measurements for the modulated barrier and deep etched quantum wire and dot structures. Results for the exciton binding energies for modulated barrier structures are shown in Fig. 4 and are seen to be in good agreement with the data. Our calculated results for the diamagnetic shifts and binding energies help to validate the method used to obtain the experimental binding energies from the diamag-

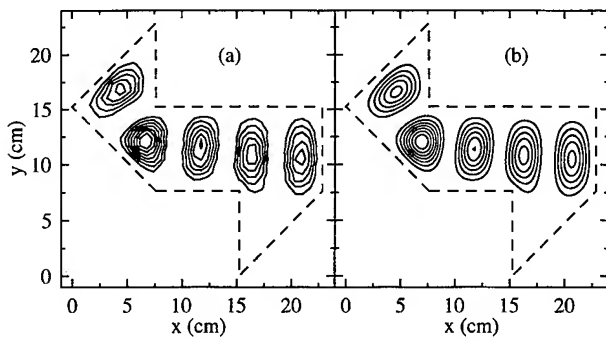


FIG. 5. (a) Measured (see Ref. 13) and (b) calculated values of E^2 for the sixth mode of a thin microwave cavity with the shape given by the dashed lines.

netic shifts. For the modulated barrier systems, the binding energies decrease for very small sizes after increasing to a maximum with decreasing size. This behavior results from the exciton wave function extending outside of the wire or dot region at small sizes. From these results, it is seen that exciton binding energies in these structures can be increased up to about 3 1/2 times the bulk value of 4.03, a result that is of interest in connection with applications in detectors and modulators. Even more surprisingly, significant enhancement of the binding occurs for sizes up to ten times the exciton Bohr radius (~ 10 nm). We have traced this behavior to the fact that center-of-mass coordinates and relative coordinates are not separable, and in these systems both are confined.

Recently, an enhancement of the exciton binding energy up to six times the bulk value in a "T-wire" structure formed at the intersection of two quantum wells each of about 5.4 nm width was reported.¹¹ We have made detailed calculations for this system using the method given above, and we find that the exciton binding energy for this structure should be about 3 1/2 times the bulk value, which is similar to the maximum enhancement seen in the barrier modulated and deep etched structures above. We suggest that the discrepancy between our calculations and the experimental value arises from shortcomings in the analysis of the luminescence data.

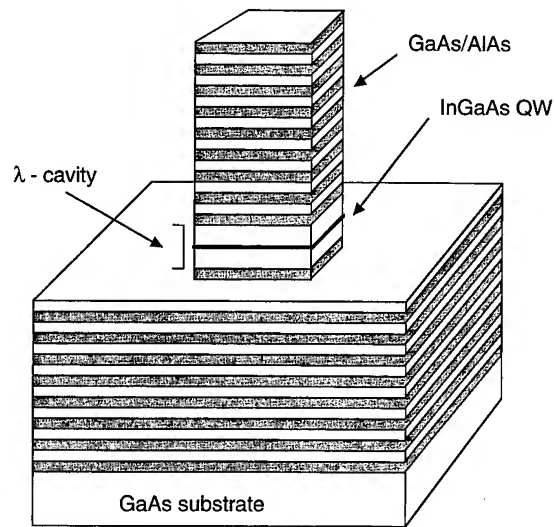
Confined photons in microcavities and waveguides have attracted much interest recently. In order to treat these problems, we have developed a boundary element method to treat the electromagnetic modes described by Maxwell's Equations.⁵ For this case, the components electric fields E_α are determined by

$$(\nabla^2 + K^2)E_\alpha(\mathbf{r}) = 0, \quad (8)$$

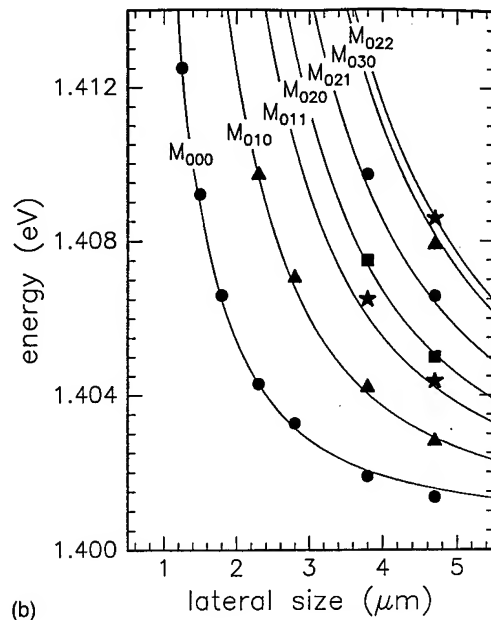
where $K^2 = \epsilon\omega^2/c^2$, and ϵ is the high-frequency dielectric constant. Here, the Green's function equation obeys

$$(\nabla^2 + K^2)G(\mathbf{r}, \mathbf{r}'; \omega) = \delta(\mathbf{r} - \mathbf{r}'), \quad (9)$$

and for a uniform region of space, it is $G(\mathbf{r}, \mathbf{r}'; \omega) = e^{iKr}/4\pi r$ for $r = |\mathbf{r} - \mathbf{r}'|$. Then, the fields are determined by equations similar to Eqs. (3) and (4). We have found that calculations of the photon states in confined structures using this method converge very well with the number of boundary elements. In particular, they appear to converge considerably



(a)



(b)

FIG. 6. (a) Sketch of the semiconductor microcavity. (b) Energies of confined photon modes of the cavity vs lateral size from photoluminescence studies. Solid lines give the results of the calculations.

more rapidly than do plane-wave expansion methods¹² for photon states, which typically require hundreds of waves. Thus, the present method should provide a useful approach for solving the photonic band-gap problem for the full vector Maxwell's Equations.

In order to illustrate these results for photons, we give two examples. The most detailed measurements to date of confined photons have been made for thin microwave cavities¹³ in shapes of interest in studies of the onset of quantum chaos. Maxwell's Equations for such cavities become isomorphic to Schrödinger's equation in two dimensions. Contour plots of the measured and calculated values of the square of the electric fields for one mode are shown Fig. 5. It is seen that these calculations are in good agreement with the experimental results.

In recent experiments, confined optical modes were studied in semiconductor microcavities with lateral sizes on the order of a few microns.¹⁴ Here, λ -wave surface emitting laser cavities with GaAs/AlAs mirrors and InGaAs quantum wells in the cavity were etched laterally. A sketch of these structures is shown in Fig. 6. Here, luminescence from the photons arises from the weak coupling of the confined photons with excitons in 7 nm InGaAs quantum wells in the center of the cavity. The measured energies of the ground state and of higher lying confined photon modes are shown in Fig. 6 with the corresponding theoretical results. It is seen that these calculations give a good description of the effects of confinement on the photons. From them, we find that the photons can give a good approximation represented by modes whose fields go to zero at the lateral boundaries of the cavity. In this way, the modes can be characterized by the number of nodes in each of their two lateral directions. From the calculations, the modes are found to extend somewhat into the vacuum and substrate regions. In addition, they couple to the continuous spectrum of modes of the vacuum, and thus, strictly speaking, they are strongly confined resonances. These results will be of considerable interest in connection with improved laser action because recombination into discrete modes can reduce or eliminate spontaneous emission.

Acknowledgment: This work was supported in part by the U.S. Office of Naval Research.

- ¹Arvind Kumer, *Surf. Sci.* **263**, 335 (1992).
- ²T. L. Reinecke and P. A. Knipp, in *Advances in Solid State Physics* [Festkörperprobleme Vieweg, Wiesbaden (1997)], Vol. 36, p. 105.
- ³P. A. Knipp and T. L. Reinecke, *Phys. Rev. B* **45**, 9091 (1992).
- ⁴P. A. Knipp and T. L. Reinecke, *Phys. Rev. B* **54**, 1880 (1996).
- ⁵P. A. Knipp and T. L. Reinecke, *Proceedings of the 23rd International Conference on the Physics of Semiconductors, Berlin, 1996*, edited by M. Scheffler and R. Zimmermann (World Scientific, London, 1996), pp. 3131.
- ⁶Ch. Gréus, F. Daiminger, A. Forchel, P. A. Knipp, and T. L. Reinecke, *Phys. Rev. B* **47**, 7626 (1993).
- ⁷Ch. Gréus, A. Forchel, P. A. Knipp, T. L. Reinecke, R. Spiegel, and F. Faller, *Phys. Rev. B* **49**, 5753 (1994).
- ⁸M. Bayer, A. Schmit, A. Forchel, F. Faller, T. L. Reinecke, P. A. Knipp, A. A. Dremin, and V. D. Kulakovskii, *Phys. Rev. Lett.* **74**, 3439 (1995).
- ⁹M. Bayer, S. Walck, T. L. Reinecke, and A. Forchel, *Europhys. Lett.* (in press).
- ¹⁰Y. Nagume, Y. Arakawa, S. Tsukamoto, and M. Nishioka, *Phys. Rev. Lett.* **69**, 2963 (1992).
- ¹¹S. Someya, H. Akiyama, and H. Sakaki, *Phys. Rev. Lett.* **76**, 2965 (1996).
- ¹²R. D. Meade, A. M. Rappe, K. D. Brommer, and J. D. Joannopoulos, *Phys. Rev. B* **48**, 8434 (1993).
- ¹³S. Sridhar and A. Kudrolli, *Phys. Rev. Lett.* **72**, 2175 (1994).
- ¹⁴J. P. Reithmaier, M. Röhner, H. Zull, F. Schäfer, A. Forchel, P. A. Knipp, and T. L. Reinecke, *Phys. Rev. Lett.* **78**, 378 (1997).

Electronic structure and optical behavior of self-assembled InAs quantum dots

K. W. Berryman,^{a)} S. A. Lyon,^{b)} and Mordechai Segev^{b),c)}

Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544

(Received 12 January 1997; accepted 3 April 1997)

Using both electrical and optical techniques, ground state energy levels and excited states of carriers in self-assembled InAs quantum dots are described, and the first observations of mid-infrared photoconductivity in these structures are presented. Electrical measurements including ac conductance and Hall techniques have been used to determine thermal trapping of carriers, and yield strong binding for holes, less binding for electrons, and an exciton energy that is consistent with photoluminescence (PL) measurements. Further PL experiments have probed the effect of changing the InAs dot size, and using embedding material of different composition. Several of these structures also demonstrate strong electron binding. In devices where the InAs dots have been grown in an $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ matrix and surrounded by AlAs barriers, normal incidence photoconductivity has been observed at a range of wavelengths in the mid-infrared and attributed to single carrier transitions out of the dots. This mid-infrared optical response is investigated for several different dot structures and compared to photoluminescence data from the same samples. © 1997 American Vacuum Society. [S0734-211X(97)07404-0]

I. INTRODUCTION

In the past several years there has been a surge of interest in structures which exhibit quantum confinement in three dimensions, commonly known as quantum dots (QDs). A number of approaches to fabricating semiconductor quantum dots have been pursued: direct lithographic patterning,^{1,2} regrowth on etched substrates,^{3,4} patterned "stressors" to induce local strain fields,^{5,6} embedding semiconductor clusters in glassy hosts,⁷ and several chemical techniques for synthesizing and encapsulating nanocrystals.^{8,9} Each of these methods has advantages and drawbacks, but none has been able to produce a high density of nearly uniform dots embedded without defects in a semiconductor host.

Recently, however, it has been discovered that dots can be "self-assembled" during the pseudomorphic growth of highly lattice-mismatched semiconductors.¹⁰⁻¹² InAs and InGaAs dots embedded in GaAs is probably the most widely studied system which exhibits this behavior, but the self-assembly of dots in a variety of mismatched heterostructures has been demonstrated.¹³ The dots arise from a modified Stransky-Krastinow growth mechanism in which about 1 monolayer of InAs uniformly covers the GaAs but the subsequent InAs aggregates into three dimensional islands strained to the GaAs because of the 7% lattice mismatch between InAs and GaAs. Atomic force microscopy shows low pyramids, with base lengths of about 100–200 Å and a height of about 20–30 Å.¹⁴ Of particular importance is the fact that subsequent growth is epitaxial, embedding the pseudomorphically strained InAs quantum dots in GaAs of high crystal quality. Furthermore, extremely high dot cover-

age ($\geq 5 \times 10^{10}/\text{cm}^2$) has been demonstrated.¹⁵ Perhaps most importantly, these self-assembled QDs have been shown to have the atomiclike transitions that are the hallmark of a discrete density of states.^{16,17}

The discrete density of states in this QD system has important consequences for the dynamics of carriers within the QDs, and therefore also for the optical properties of these dots. In particular, since the state spectrum is discrete, LO phonon scattering should be strongly inhibited because of a lack of available states, which implies long carrier lifetime and dephasing times. Also, since the QDs are embedded in the lattice without defects, they would be expected to exhibit a high radiative efficiency. In fact, for InAs dots in a GaAs matrix successful laser operation has already been achieved in the near-infrared with QDs as the gain material.¹⁸ Other workers have shown that even for InAs QDs grown on a GaAs-on-Si substrate, which has substantial defects, the radiative efficiency of the dots is high.¹⁹ All of these results, however, were obtained in the near-infrared on the exciton transition. To the best of our knowledge, other than a couple of experiments reporting far-infrared absorption in charged InGaAs QDs,²⁰ and recently far-infrared absorption in charged InAs QDs,²¹ there has been little work on transitions involving a single carrier. In this work we discuss measurements of the ground state energies of single carriers in InAs QDs in a variety of structures of different design, then present the first observations of mid-infrared photoconductivity in this system. We discuss the photoconductivity observed in several structures, and compare it with photoluminescence (PL) and electroluminescence (EL) performed on the same devices. We believe this work provides a foundation for the development of self-assembled QDs as efficient mid-infrared detectors and emitters.

^{a)}Electronic mail: berryman@ee.princeton.edu

^{b)}Members of the Advanced Technology Center for Photonics and Optoelectronic Materials (POEM).

^{c)}Member of the Princeton Materials Institute (PMI).

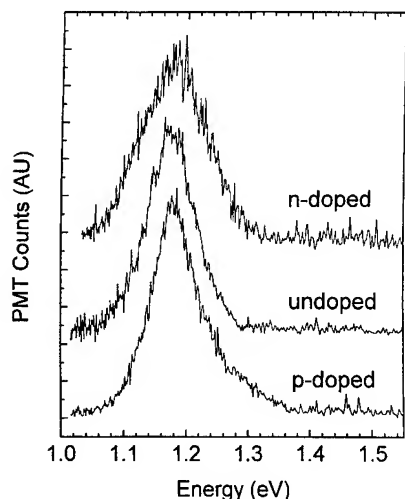


FIG. 1. Photoluminescence spectra at 80 K of *n*-doped (top curve), undoped (middle curve), and *p*-doped (bottom curve) InAs QDs. The curves have been offset for clarity.

II. ENERGY LEVELS IN INAS QUANTUM DOTS

Before optical experiments can be performed on a new material system, it is necessary to determine at which wavelengths the system should respond. In a QD system, the optical response from individual carriers can result either from transitions between energy levels in the QD, or ionization of the carrier out of a particular energy level. In either case, knowing the position of the carrier energy levels, and particularly of the ground state, is of primary importance. The ground state energy of an exciton in a QD is readily measured by PL experiments, with typical results indicated in Fig. 1. This figure presents PL spectra acquired for three samples: undoped QDs, dots doped with electrons, and dots doped with holes. The doping density used, 5×10^{10} impurities/cm² on the dot layer, should result in approximately 1 carrier per dot. The doping was confined to the dot layers themselves. The PL spectra were acquired with an excitation intensity of approximately 0.2 W/cm² at 632.8 nm, with the samples held at 80 K in a nitrogen cryostat. The reproducibility of the PL peak positions between these different samples implies that QDs with similar size distributions have been formed on each sample, and that the dopants have not affected the energy level structure. These samples were all prepared using standard molecular beam epitaxy (MBE) growth techniques on GaAs(001) substrates. After depositing a 0.5 μm buffer layer of GaAs, the dots were fabricated by lowering the substrate temperature to 500 °C and growing slightly over 2 monolayers of InAs. The transition to three dimensional growth was confirmed by changes in reflection high-energy electron diffraction (RHEED) patterns. After allowing the dots a short interval to reach equilibrium, during which time dopants were introduced in the case of doped samples, the substrate temperature was raised to 610 °C and a cap layer, typically 0.3 μm thick, was grown on top of the dots. All of the samples discussed here were prepared in similar fashion.

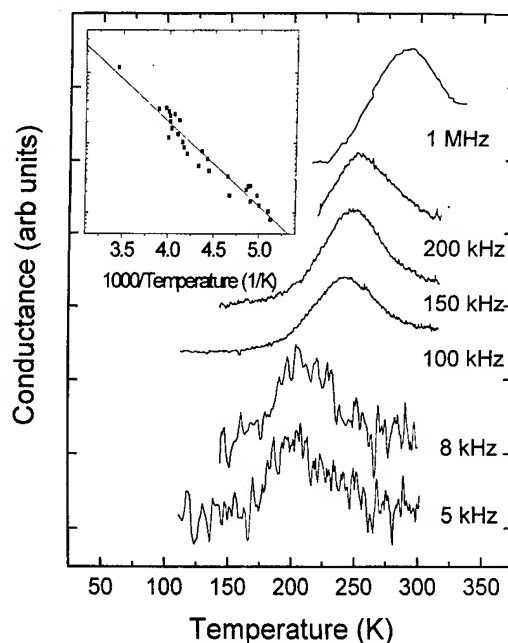


FIG. 2. ac conductance as a function of temperature. Each curve corresponds to a different measurement frequency, as indicated. The inset is an Arrhenius plot of T^2/f vs $1/T$, where T is the temperature of the conductance peak which occurs at measurement frequency f . The straight line fit in the inset indicates a thermal trapping depth for holes in these quantum dots of 240 meV.

The PL results in Fig. 1 directly indicate an exciton ground state binding energy, given by the difference between the observed peak position and the GaAs band edge, of ~ 300 meV for the QDs observed in Fig. 1. Unfortunately, PL provides no indication of how this binding energy is shared between electrons and holes. So to resolve the binding of individual carriers, we began a series of electrical measurements on several QD samples. For one set of experiments, we placed QDs in a *p-n* junction and measured the temperature dependence of the ac conductance as a function of frequency. In another set of experiments, using the Hall effect we measured the carrier concentration in QD layers as a function of temperature. Both of these techniques yield the thermal trapping depth of carriers in QDs, which can be determined independently for *n*- and *p*-doped samples, thereby providing a direct measure of electron and hole binding.

Although the full details of these measurements are presented elsewhere,²² an example of the ac conductance data is indicated in Fig. 2. This figure shows the set of conductance peaks in temperature at different frequencies which result when the thermal emission rate from the dots matches the ac measurement frequency. In this case, the QD sample was *p* doped. The inset shows an Arrhenius plot of T^2/f vs $1/T$, where T is the temperature of the conductance peak at a measurement frequency f . The linear fit shown in the inset yields a thermal trapping depth for holes of 240 meV. Graphs of this sort, and indeed the entire method of capacitance spectroscopy, have long been used to determine the depth of defect traps in semiconductors.²³ In this case the traps are the quantum dots themselves. Temperature depen-

dent Hall measurements provide an independent measure of the thermal trapping depth. In this case the change in carrier concentration as a function of temperature is assumed due to trapping in dots; fitting the temperature dependence also yields a trap depth for holes of ~ 240 meV, and indicates a binding energy for electrons of a few tens of meV. The total binding is therefore in reasonable agreement with that obtained from PL observations.

We conclude that for the size distribution of QDs measured here, holes are strongly bound while electrons exhibit much less binding. These results are in agreement with several recent theoretical calculations^{24–26} and experiments.^{27,28} One note of caution must be observed here: the energy levels for QDs of different size will vary, so comparisons can be difficult without some measure of size. Fortunately, recent measurements comparing PL emission energies with QD size determined by transmission electron microscopy (TEM) have shown that the PL emission peak correlates well with the dot size.²⁹ For the QDs measured using the ac conductance and Hall techniques, the position of the PL peak then yields an estimated base size of 10 nm, for which the calculated energy levels in Ref. 24 are 200 meV for holes and 90 meV for electrons—in reasonable agreement with the experimental values.

III. QUANTUM DOT ENGINEERING

Since optical devices based on electronic transitions are both easier to design and understand than those involving holes, we chose to attempt to manipulate the energy levels in the QD system to increase the electron binding. Our first attempt involved simply changing the matrix surrounding the QDs from pure GaAs to $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$. At this Al concentration, the band-gap of the AlGaAs material is about 1.9 eV, as compared to about 1.5 eV for pure GaAs. As the wave functions of electrons and holes inside a QD are largely confined to the InAs, we expected that the absolute energies of electrons and holes would change little, which would imply that much of the 400 meV additional binding could be present for the electronic states.

We prepared two samples which were identical except that one consisted of dots grown in GaAs, while for the other we first grew a layer of $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$, then grew QDs, then overcoated with $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ as well. In Fig. 3 we see PL spectra which confirm our initial suspicion—the ground-state energy of the QDs in AlGaAs has moved up by only 100 meV. Since there is about an additional 250 meV of conduction band offset in the AlGaAs as compared to GaAs, the electrons in the QDs in AlGaAs must be bound by at least an additional 150 meV.

In an attempt to produce even greater electron binding, we designed samples with a variety of cladding structures surrounding the QDs. We used a combination of 20-Å-thick AlAs barriers and 10 Å GaAs layers on top of the $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ matrix to change the confining potential seen by the InAs dots. The dot structures are illustrated in Fig. 4 and were separated during growth by 1000 Å of undoped $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$, both to prevent any coupling between layers

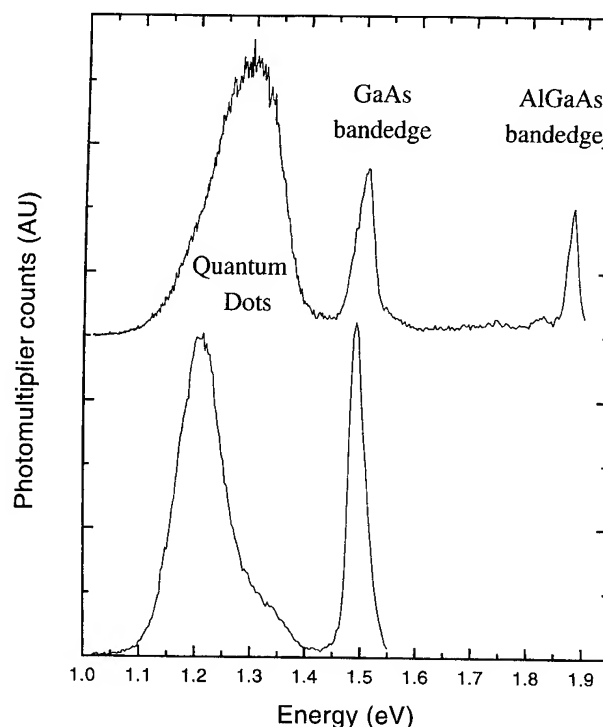


FIG. 3. Photoluminescence spectra at 80 K of standard InAs dots in GaAs (bottom curve), and from identical dots in $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ (top curve). The GaAs band edge (near 1.5 eV), and $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ band edge (near 1.9 eV) are also evident.

and also so that different dot layers could be exposed by chemical etching. In these samples the dots were also lightly n doped with a surface density of 5×10^{10} carriers/cm², resulting in about 1 electron per dot.

The samples were then characterized at a temperature of 80 K with PL measurements. Two typical PL spectra using 488 nm light at an intensity of 25 W/cm² are illustrated in Fig. 5. These spectra were acquired with a nitrogen cooled photomultiplier tube (PMT) system which is losing sensitivity rapidly at the lowest energies shown. The bottom curve includes emission from all three QD layers, while the top curve was taken after the layer containing structure A had been etched away using a sulfuric acid based etch. The ratio of peak intensities in structures B and C has changed because structure C is now closer to the surface, and therefore sees a

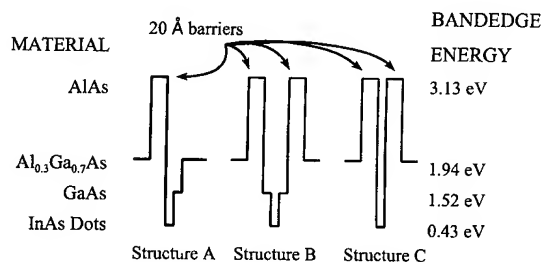


FIG. 4. Schematic drawing of the three QD structures, labeled A, B, and C, used for photoluminescence and photoconductivity measurements.

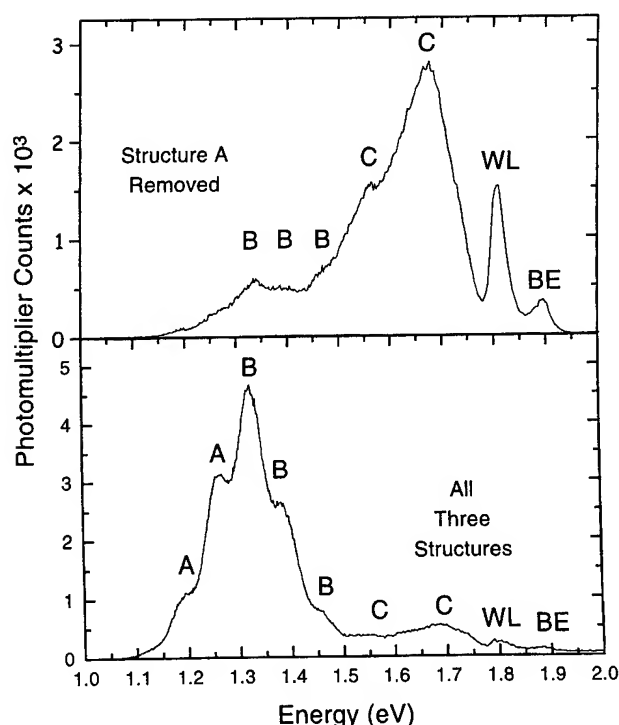


FIG. 5. The bottom curve is a photoluminescence spectrum at 80 K of a sample containing the three QD structures illustrated in Fig. 4. The top curve is a PL spectrum in which structure A has been etched away. The peak assignments indicated in the figure were determined by removing specific layers by etching and observing resultant changes in the peak amplitudes. Photoluminescence from the wetting layer (WL) and band edge (BE) are also observed.

higher density of electron hole pairs created by the incident laser. Structure A emission has almost entirely disappeared, although some remnant emission may still be present because the etch is not perfectly uniform. By performing a sequence of such etching operations and measuring the changes in the PL peaks, we were able to assign the emission to specific dot layers as indicated in the figure. Furthermore, the multiple peaks for individual QD layers result from emission by excited states, as the ground states are saturated at this excitation intensity.³⁰ Note that this is 75 times the intensity used in Fig. 1, which showed no evidence of excited state emission. We should also note that the subpeak structure in Fig. 5 is not due to interference effects, both because this structure does change with decreasing excitation intensity, and because the spacing implied by the peaks is inconsistent with the growth thickness.

The data in Fig. 5 illustrate that by employing suitable barrier structures we have great freedom in manipulating the ground state energy of carriers in the QDs. In structure A, for example, there is almost 700 meV of binding in the exciton, implying that both carriers likely have several bound states. In structure C, by contrast, the exciton ground state binding is less than 300 meV, implying that at least one carrier must have its ground state in the QD above the corresponding GaAs band edge. This in turn suggests that QDs of this form

might be useful in resonant tunneling devices, including quantum cascade lasers.³¹

IV. PHOTOCONDUCTIVITY IN INAS QUANTUM DOTS

Since we were now confident that our cladded QD samples contained electrons which were strongly bound in the dots, we began a series of experiments to attempt to observe optical transitions by single carriers within the QDs. We chose to look for photoconductivity, using Schottky gates on samples containing dots to allow us to see a photo-voltaic optical response. We used the structures, illustrated in Fig. 4, which had been lightly n doped for this purpose, and etched samples to different depths to expose specific QD layers. We evaporated semi-transparent 100 Å Ti gates on each sample over an area of several square millimeters, followed by a small gold pad for bonding. Back contact was made through the heavily doped substrate. Each sample was mounted in a nitrogen cryostat, then positioned at normal incidence at the focus of the output beam of a Bomem DA3 Fourier transform infrared (FTIR). A standard globar source and CaF₂ beamsplitter were used. In this configuration the FTIR could be scanned while monitoring the photoconductivity signal from the QDs, producing an interferogram which was then Fourier transformed to reveal the QD spectral response.

Although more extensive analysis and discussion of these experiments will be presented elsewhere,³² two typical optical responses at different applied biases are illustrated in Fig. 6. At short wavelengths, a broad Schottky response is observed, identical to the response from control samples without QDs. The control samples, however, show no changes with bias and no response in the mid-infrared, while the QDs have a large response which changes with applied bias. We believe the QD response is due to electrons being excited from the QD states into the Al_{0.3}Ga_{0.7}As conduction band, where they are swept out of the QDs by the internal electric field, resulting in current flow in the external circuit. A response of this sort requires that QDs with trapped carriers be present in the vicinity of an internal electric field. Such a situation occurs at the edge of the Schottky barrier depletion region.

A consequence of this explanation is that we would expect to be able to select which dot layer responds by changing the dc bias applied to the Schottky contact—which is exactly what we observe. As the forward voltage is increased (decreased), the QDs which are shallower (deeper) in the device will respond. In fact, in Fig. 6 we see that the primary response peak shifts from 390 to 300 meV with a change in applied bias of about 1 V. This energy difference compares favorably with the 100 meV separation between structure A and structure B exciton ground states observed in Fig. 5. It is important to note here that the energy separation observed in the exciton ground states must be larger than that observed in photoconductivity, because the photoconductivity in an n-doped sample will not respond to changes in the position of the hole energy levels, which do nonetheless affect the

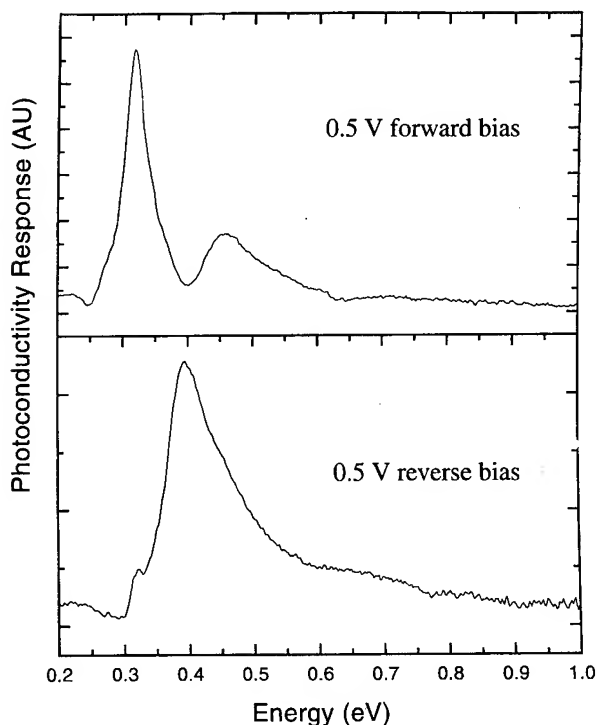


FIG. 6. Photoconductivity spectra at two different applied biases for the sample containing the three structures in Fig. 4. As the bias is changed, different populations of dots contribute to the response as the edge of the depletion region changes position. The curves have been offset vertically for clarity.

exciton energy. The fact that almost all of the shift is accounted for in the photoconductivity, however, implies that for these structures, the AlAs barriers have the greatest impact on the electron energy levels. This is to be expected as the electronic wave functions, because of the light electronic mass, should extend further into the barriers than the holes.

V. DISCUSSION AND CONCLUSIONS

Self-assembled QDs represent perhaps the first system which combines the power and flexibility of modern MBE growth techniques with the scientific and technological interest of three dimensionally quantum confined structures. We have seen here that by proper design the individual energy levels of electrons and holes within these dots can be tailored over a wide range, and can be measured using standard electronic and optical techniques.

Using structures optimized to produce large electron binding, to the best of our knowledge the measurements presented here are the first observations of mid-infrared photoconductivity in self-assembled dots. This response is not present in samples without QDs, and disappears from samples in which the QDs have been etched away. The large apparent optical response and the many wavelengths accessible by varying structure design argues persuasively that QDs might be useful not only as infrared detectors, but also possibly as emitters. QDs detectors could expect to exhibit lower dark noise than is present in a quantum well detector, which is the dominant source of noise in these devices.^{33,34}

Furthermore, the reduced phonon scattering which results from the discrete density of states in a QD should result in long lifetime and dephasing times, increasing their radiative efficiency and perhaps allowing novel coherent nonlinear optical experiments.³⁵ The results presented here for these QD structures should therefore serve as the starting point for further exploration of the electronic and optical properties of these artificial atoms.

ACKNOWLEDGMENTS

This work was supported in part by the MRSEC program of the National Science Foundation under Award No. DMR-9400362, and the URI program of the Army Research Office under Grant No. DAAL03-92-G-0146.

¹J. N. Randall, M. A. Reed, and Y. -C. Kao, *J. Vac. Sci. Technol. B* **8**, 1348 (1990).

²A. Scherer and B. P. Von der Gaag, *Proc. SPIE* **1284**, 149 (1990).

³H. Sakaki, *Curr. Opin. Solid State Mater. Sci.* **1**, 17 (1996).

⁴E. Kapon, G. Biasiol, D. M. Hwang, M. Walther, and E. Colas, *Solid-State Electron.* **40**, 815 (1996).

⁵K. Kash, R. Bhatt, D. D. Mahoney, J. M. Worlock, P. S. D. Lin, A. Scherer, B. P. Von der Gaag, M. Koza, and P. Grabbe, *Advances in Materials, Processing and Devices in III-V Compound Semiconductors Symposium* (Boston, MA, 1989), p. 245.

⁶K. Kash, D. D. Mahoney, B. P. Von der Gaag, A. S. Godz, J. P. Harbison, and T. Florez, *J. Vac. Sci. Technol. B* **10**, 2030 (1992).

⁷M. G. Bawendi, W. L. Wilson, L. Rothberg, P. J. Carroll, T. M. Jedjv, M. L. Steigerwald, and L. E. Brus, *Phys. Rev. Lett.* **65**, 1623 (1990).

⁸A. P. Alivisatos, *Science* **271**, 933 (1996).

⁹J. A. Lebens, C. S. Tsai, K. J. Vahala, and T. F. Kuech, *Appl. Phys. Lett.* **56**, 2642 (1990).

¹⁰F. Houzay, C. Guille, J. M. Moison, P. Henoc, and F. Barthe, *J. Cryst. Growth* **81**, 67 (1987).

¹¹M. Tabuchi, S. Noda, and A. Sasaki, *Science and Technology of Mesoscopic Structures*, edited by S. Namba, C. Hamaguchi, and T. Ando (Springer, Tokyo, 1992), p. 379.

¹²M. Tsuchiya, J. M. Gaines, R. H. Yan, R. J. Simes, P. O. Holtz, L. A. Coldren, and P. M. Petroff, *Phys. Rev. Lett.* **62**, 466 (1989).

¹³E. R. Glaser, B. R. Bennett, B. V. Shanabrook, and R. Magno, *Appl. Phys. Lett.* **68**, 3614 (1996).

¹⁴J. M. Moison, F. Houzay, F. Barthe, L. Leprince, E. Andre, and O. Vatel, *Appl. Phys. Lett.* **64**, 196 (1994).

¹⁵G. S. Solomon, J. A. Trezza, and J. S. Harris, *Appl. Phys. Lett.* **66**, 3161 (1995).

¹⁶J. -Y. Marzin, J. -M. Gerard, A. Izrael, D. Barrier, and G. Bastard, *Phys. Rev. Lett.* **73**, 716 (1994).

¹⁷M. Grundmann, J. Christen, N. N. Ledentsov, J. Bohrer, D. Bimberg, S. S. Ruvimov, P. Werner, U. Richter, U. Gosele, J. Heydenreich, V. M. Ustinov, A. Y. Egorov, A. E. Zhukov, P. S. Kop'ev, and Zh. I. Alferov, *Phys. Rev. Lett.* **74**, 4043 (1995).

¹⁸D. Bimberg, N. N. Ledentsov, M. Grundmann, N. Kirstaedter, O. G. Schmidt, M. H. Mao, V. M. Ustinov, A. Y. Egorov, A. E. Zhukov, P. S. Kop'ev, Zh. I. Alferov, S. S. Ruvimov, U. Gosele, and J. Heydenreich, *Jpn. J. Appl. Phys.* **35**, 1311 (1996).

¹⁹J. M. Gerard, O. Cabrol, and B. Sermage, *Appl. Phys. Lett.* **68**, 3123 (1996).

²⁰H. Drexler, D. Leonard, W. Hansen, J. P. Kotthaus, and P. M. Petroff, *Phys. Rev. Lett.* **73**, 2252 (1994).

²¹M. Fricke, A. Lorke, J. P. Kotthaus, G. Medeiros-Ribeiro, and P. M. Petroff, *Europhys. Lett.* **36**, 197 (1996).

²²K. W. Berryman, S. A. Lyon, and Mordechai Segev (to be published).

²³E. H. Nicollian and J. R. Brews, *MOS (Metal Oxide Semiconductor) Physics and Technology* (Wiley, New York, 1982), pp. 176-234.

²⁴M. Grundmann, O. Stier, and D. Bimberg, *Phys. Rev. B* **52**, 11969 (1995).

- ²⁵M. A. Cusack, P. R. Briddon, and M. Jaros, *Phys. Rev. B* **54**, R2300 (1996).
- ²⁶A. Wojs and P. Hawrylak, *Phys. Rev. B* **53**, 10841 (1996).
- ²⁷M. Grundmann, N. N. Ledentsov, O. Stier, J. Bohrer, D. Bimberg, V. M. Ustinov, P. S. Kop'ev, and Zh. I. Alferov, *Phys. Rev. B* **53**, R10509 (1996).
- ²⁸P. N. Brunkov, S. G. Konnikov, V. M. Ustinov, A. E. Zhukov, A. Y. Egorov, M. V. Maksimov, N. N. Ledentsov, and P. S. Kop'ev, *Semiconductors* **30**, 492 (1996).
- ²⁹A. Polimeni, A. Patane, M. Capizzi, F. Martelli, L. Nasi, and G. Salviati, *Phys. Rev. B* **53**, R4213 (1996).
- ³⁰M. Grundmann, N. N. Ledentsov, O. Stier, J. Bohrer, D. Bimberg, V. M. Ustinov, P. S. Kop'ev, and Zh. I. Alferov, *Phys. Rev. B* **53**, R10509 (1996).
- ³¹J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchison, and A. Y. Cho, *Science* **264**, 553 (1994).
- ³²K. W. Berryman, S. A. Lyon, and Mordechai Segev, *Appl. Phys. Lett.* **70**, 1861 (1997).
- ³³K. W. Goossen and S. A. Lyon, *J. Appl. Phys.* **63**, 5149 (1988).
- ³⁴M. A. Kinch and A. Yariv, *Appl. Phys. Lett.* **55**, 2093 (1989).
- ³⁵Including, for example, the field-induced symmetry breaking described by A. Levinson, M. Segev, G. Almog, and A. Yoris, *Phys. Rev. A* **49**, R661 (1994).

Surface morphology and quantum dot self-assembly in growth of strained-layer semiconducting films

K. E. Khor^{a)} and S. Das Sarma

Department of Physics, University of Maryland, College Park, Maryland 20742-4111

(Received 12 January 1997; accepted 16 April 1997)

The energetics of island growth on thin semiconducting films deposited on lattice-mismatched substrates is discussed in this article. Hut clusters similar to those proposed by Mo *et al.* [Phys. Rev. Lett. **65**, 1020 (1990)] for Ge/Si(001), with $(10n)$, $(n \geq 3)$, or $(11n)$, $(n \geq 1)$, side facets will be analyzed. Results show that hut clusters are the energetically favorable structures during early growth, with side facets of rebonded (105) planes; at later times, larger islands with $(11n)$ -like facets become favorable. It is found that islands nucleate with critical nuclei of about 1 atom and grow two dimensionally until they reach a critical size s_c , when it is favorable for the islands to become three dimensional. There is an effective barrier at the transition from two dimensional to three dimensional growth. Beyond the barrier, there is an immediate energy gain which can be large, on the order of 5–10 meV/atom for the highly mismatched system of InAs/GaAs. It is suggested that these results are the underlying reason for the uniformity seen in self-assembled quantum dots in highly mismatched heteroepitaxy. © 1997 American Vacuum Society. [S0734-211X(97)08204-8]

In this work on the growth of strained-layer superlattices, we focus on the early stages for systems which grow in the Stranski–Krastanov (SK) mode. The nature of islands seen have been characterized by Mo *et al.*¹ for Ge grown on Si(001); “hut” clusters are the first type of islands to appear with well defined (105) facets, tilted at 11.3° to the surface, then there is a transition to larger islands with $(11n)$ -like faces² and finally to even larger but dislocated islands. This sequence of islanding has probably been seen in other systems, like the heteroepitaxy of InP on GaAs(001), where the mismatch is also about 4%.³ It is possible to bypass the hut cluster stage by growing at slightly higher temperatures.^{1,3} There has also been much interest on growth in even more highly mismatched systems ($\sim 7\%$) like InAs/GaAs(001), because of the uniformity in the size of islands seen.^{4,5} This uniformity, with dispersions of 10% in height and 7% in diameter of the islands at the initial stages of formation, decreases with coverage θ .² There seems to be a distinct coverage θ_c [$=1.5$ monolayers (ML),⁴ 1.75 ML,⁵ 1.7 ML⁶] at which the transition from two dimensional (2D) to three dimensional (3D) growth occurs for the InAs/GaAs system. This critical thickness transition is dependent on growth conditions; the work of Gerard *et al.*,⁶ shows that by substantially increasing the deposition rate, for example, it is possible to shift it from 1.7 to 1.95 ML. There is evidence that the material to build an island comes mainly by depleting its immediate environment: the thickness of the InAs layer before islanding occurs, which is between 1 and 2 ML and is reduced to 1 ML in the region surrounding the island.^{4–6} This suggests that the critical layer thickness for InAs/GaAs(001), beyond which it is energetically favorable to form islands, is actually 1 ML and that the extra thickness before islanding is seen may be due to the presence of a barrier at the 2D–3D transition. There is sign that some depletion is also present in the growth of Ge on Si.⁷

In this study we look at the energetics of the 2–3D tran-

sition in detail by means of molecular dynamics simulation. We use an empirical potential that has been appropriately tuned.⁸ We observed a generic trend, applicable to the materials above and not sensitive to potential parameters; because taller islands relax more, hut clusters with $(10n)$ side facets become energetically favorable at sizes which increase with n [$n=3$ is excluded as (103) facets entail effectively, D_A steps which are known to be energetically very costly⁹]. (111) faceted islands, which are taller than $(10n)$ islands, become energetically favorable at sizes bigger than the size of the latter type, especially at low n values. With this trend in mind, we show that the 2D–3D transition occurs when 2D islands have grown much larger than the size 3D islands first become energetically favorable; this effectively is a barrier, which once scaled by a 2D island, allows it to reorganize itself into a 3D shape, with an immediate gain of energy. This gain, which is more pronounced for the highly mismatched InAs/GaAs system than for the Ge/Si system, can be quite substantial, from 5 to 10 meV/atom for the former. We feel that this is the underlying factor for the uniformity of sizes of islands seen in this system. Priester *et al.*¹⁰ have attempted to provide an explanation for the uniformity of the 3D islands, but have not taken into account the factor of the barrier, which should affect their considerations.

We use an empirical potential that has been tuned to Si(001) surface energies and stresses⁸ and adjust it further to the energy of the D_B step as given by Chadi or Oshiyama.^{9,11,12} As in a previous study,⁸ simulations are carried out for 2D island clusters of height h , placed on a three-layer Ge on Si (Ge_3Si) substrate. Mo *et al.*^{1,13} have suggested the atomic configuration of (105) side facets. We show in Fig. 1 an alternative configuration where many of the dangling bonds of the former have been eliminated by rebonding at the cost of some surface strain. This rebonded side facet should also be consistent with the scanning tunneling microscope (STM) images of Mo *et al.* and Iwawaki *et al.*^{1,13} In

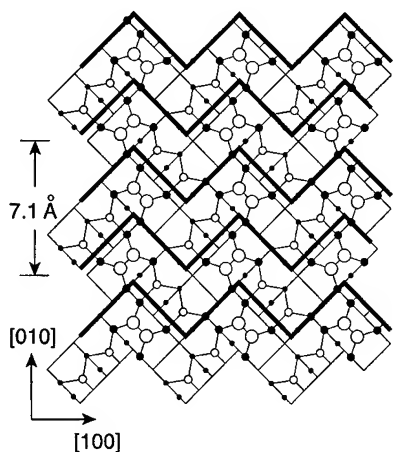


FIG. 1. Atomic configurations of rebonded (105) side facets of hut clusters. Heavy lines delineate terraces.

Fig. 2 we show the energies of these islands as a function of the size \sqrt{V} , where V is the number of atoms in an island. Hut clusters with unrebonded side facets are much higher in energy. In Fig. 2, we plot the energy of islands with (105) and (107) facets. Taller (105) faceted islands become energetically favorable at smaller sizes than (107) type islands; our calculations also indicate that the trend continues, $(10n)$ islands become favorable at sizes that increase with n . (103) faceted islands are excluded as these islands entail continuous D_A steps which are known to be substantially higher in energy than S_A , S_B , and D_B steps.⁹ This is why such islands or facets (with angles of 18.4°) are not seen experimentally. This trend of flatter islands becoming favorable at larger sizes should be generally applicable to other systems, where strain driven islanding is seen. Under certain constraints, which are reasonable for low hut clusters, this result can also be obtained from the energy expression for the elastic relaxation of an island.^{8,14} In Fig. 2, the curves for islands with small h/s (like the $(10n)$ islands), were obtained

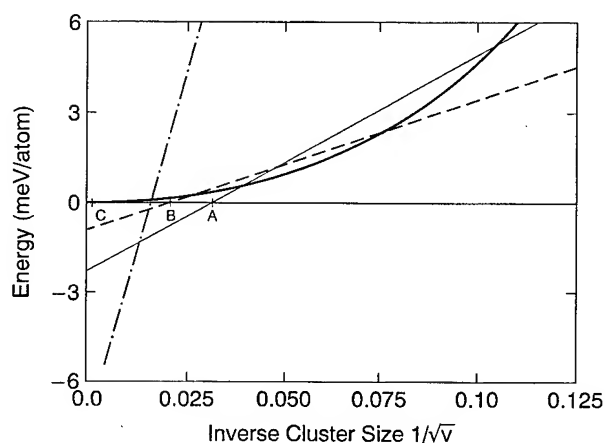


FIG. 2. Island energies in meV/atom vs cluster size. Dashed line: rebonded (105) islands; light solid: (107) islands; dot-dashed line: (111) islands; heavy solid line: two layer islands. (105), (107) and two-layer island curves cross the abscissa at points A, B, and C, respectively.

for those with minimum sized (2–4 atomic spacings) top terraces. We can show,^{8,12,14} that for such islands, increasing the size of the top terrace would increase the corresponding critical island size V_c (at which the island becomes favorable). This means that we cannot consider island energetics by imposing the constraint of a fixed specific shape on all islands. For shapes like truncated spheres, for example taller sharper islands with better energetics, are missed out. In Fig. 2 we complete the trend of (105), (107) ... clusters with a plot of the energy of two layer islands whose edges consist of a double step of the D_B type and whose top has essentially the same missing dimer (MD) configuration as the Ge_3Si surface. V_c for this island is greater than the critical sizes of (105), (107) ... faceted islands. It is known for the growth of $\text{Si}/\text{Si}(001)$,^{15–17} that islands nucleate with a critical size of 1–3 atoms and then grow two dimensionally. Island nucleation of Ge on $\text{Si}(001)$ should be similar (in both cases the dimer is the stable nucleus). We suggest that the 2D–3D transition picture is the following: 2D islands nucleate with critical nuclei of about 1 atom and grow two dimensionally until a critical size s_c , when strain makes it favorable for there to be a transition to 3D growth. This size s_c is quite large, with dimensions \sim a few hundred angstroms, as we shall see later. There is direct experimental evidence for this picture of growth. Mo and Lagally¹⁸ observe, after growth of about 3ML of Ge on Si at 500°C , a growth front roughness of three layers over an area of $60\text{ nm} \times 60\text{ nm}$. Gerard *et al.*⁶ observe one layer roughness over extensive 2D areas ($\sim 2000\text{ \AA}$) for the growth of InAs on $\text{GaAs}(001)$ at 520°C . We stress that $s_c \neq \sqrt{V_c}$ for (105) or (107) clusters, which at these stages (A or B in Fig. 2) have heights of more than 12 layers, which is much greater than the roughness over the 2D regions. Therefore the 2D island must reach a size comparable to the V_c for the two-layer island (C in Fig. 2) before it becomes energetically favorable. Once this size barrier is reached, the transition to islands of two or more layers in height is possible since taller $(10n)$ islands are already favorable at smaller sizes. There is a rapid rearrangement of its atoms in order to achieve the shape of the optimally energetic (105) faceted clusters. There is an immediate gain in energy of 1–2 meV/atom for the Ge/Si system; for InAs/GaAs we estimate this gain, assuming that the elastic energy scales with the square of misfit, to be 5–10 meV/atom. This latter amount is substantial and is probably the reason for the phenomenon of depletion seen in the highly mismatched systems. In Fig. 2 we show the results for (111) faceted islands from a previous work.⁸ We expect $(11n)$ faceted islands to show the same trend in energy; they should become favorable at sizes which increase with n ; we can also see that these sizes are at least larger than V_c for (105) islands and possibly a few other $(10n)$ islands as well.

We now focus on the experimental results of islanding in InAs/GaAs systems because of the availability of data at small increments of coverage θ .^{4–6} In Table I we put together some of the experimental results of the three groups. Growth temperatures were from 500 to 530°C and in this small range, we see that island sizes are larger with higher

TABLE I. Growth conditions and experimental results in InAs/GaAs systems.

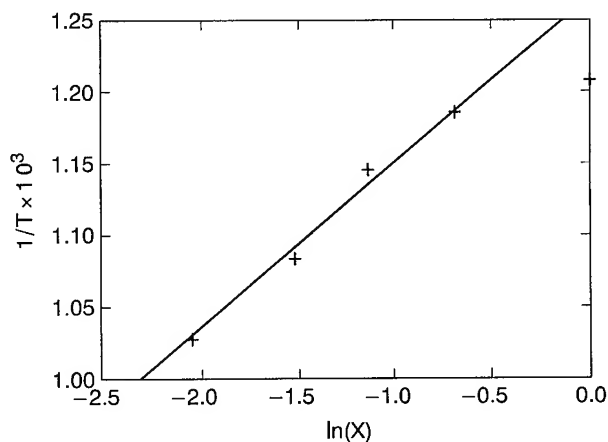
	Moison ^a	Gerard ^b	Leonard ^c
Growth temperature	500 °C	520 °C	530 °C
Deposition rate in s/ML(θ_c)	16(1.75)	(a)1–16(1.7); (b)<1(1.95)	100(1.5)
	Island features		
height;base	30 Å;240 Å		60–70 Å;200 Å
uniformity:height;base	20%; 10%		10%; 7%
notes	(410),(511) island facets	depletion zone ~1000Å	uniformity best initially island density but not size increase with θ

^aReference 5.^bReference 6.^cReference 4.

growth temperatures. This has also been observed elsewhere for fixed deposition rates.¹⁹ The smaller islands seen by Moison *et al.*⁵ have aspect ratios (height/base) consistent with (104) facets (which are inclined at 14° to the surface), while those of Leonard *et al.*⁴ have much higher aspect ratios (corresponding to facet angles $\geq 35^\circ$). These two aspect ratios are respectively consistent with (105) and (11 \bar{n}) geometries. It has been observed by Mo *et al.* that hut clusters can form at ~ 800 K, but by just increasing growth temperatures to 850 K only macroscopic islands are seen.¹ Similar results have been obtained by Sopanen *et al.*³ for the growth of InP/GaAs(001). Deposition rates must affect the layer thickness θ_c at which the 2D–3D transition occurs; at very high rates, < 1 s, the change can be as much as .25 ML in θ_c .⁴ There is uniformity in the size distributions of the islands, which has dispersions of 10% in height and 7% in diameter at the first appearance of the islands, at $\theta \sim \theta_c$; with further deposition, this uniformity is reduced, island density increases, but sizes remain essentially the same.^{4,20} All three groups concluded that there is depletionlike behavior, the coverage just before islanding is θ_c , but reducing to 1 ML below each island after it appears. Leonard *et al.* show that more than 80% of the atoms which form an island come from its environment, rather than from additional deposition. Gerard *et al.* display an atomic force micrograph (Fig. 3 in Ref. 6) of the depletion zone around an island, whose size is ~ 1000 Å. They also show that the time scale of this mass movement to form an island is from 2 to 10 s. This phenomenon of depletion is clearly consistent with the results of our simulation. Since it takes a few seconds in a highly mismatched system, we expect this effect to take much longer in the Ge/Si system; it is probably masked by the deposition rates used and only one group has reported seeing it in this system.⁷ Figure 2 shows that once islanding begins at size s_c , (point C), the immediate energy gain to form (105) islands is ~ 2 meV/atom for Ge/Si (5–10 meV/atom for InAs/GaAs) and ~ 6 meV/atom for (111) islands (~ 16 meV/atom for InAs/GaAs). As (11 \bar{n}) islands are formed at higher temperatures, this may mean that they face a higher barrier in that s_c is larger for direct transition to (11 \bar{n}) than to (105) islands. Beyond critical coverage θ_c , as deposition continues, there is much more energy to be gained for the new material to create new islands than to grow existing ones, so

there is an increase of island density but little size gain. Solomon *et al.*²⁰ have shown that 3D island density, at fixed coverage and temperature, can be increased by reducing either the growth rate R or the flux V/III ratio; the latter effectively increases the surface diffusion D . This is, at first sight, unexpected for it is known from nucleation theory for regular island growth that island density goes as R^p/D^q , where p and q are positive,¹⁷ but this applies in our case to the 2D islands. Increasing deposition rate or decreasing diffusion then increases the 2D island density and correspondingly decreases 2D island size at a given coverage, θ . This means that the first 2D islands reach size s_c at a higher critical coverage θ_c . The density of 3D islands then increases with $\Delta\theta = \theta - \theta_c$ which, as we see in the experiment above, can be attained by reducing growth rate or increasing D .

The highest uniformity in the size of 3D islands is seen at the beginning of islanding, termed “primary” islands by Leonard *et al.*⁴ Island uniformity at higher coverages is simply a consequence of this and the fact that islands remain constant in size once they are formed. The uniformity of primary 3D islands is related to the fluctuations of shape, height, and presence of defects in 2D islands approaching critical size s_c . “Trench”-like defects^{21,22} are seen on Ge₃Si surfaces, but their contributions to island size distribution are shown to

FIG. 3. Replot of the film morphology curve of Bean *et al.* (Fig. 1 in Ref. 26), $1/T$, in $K^{-1} \times 1000$, vs $\ln(x)$.

be small.¹⁵ Mo *et al.* report height fluctuations of two layers in 2D islands at temperatures of about 500 °C, so we assume that this will not be an important factor. Shape fluctuations in 2D islands are essentially step fluctuations; the entropic contributions to the free energy of 2D islands can be discussed in terms of fluctuations of steps, which would result in the smearing out of the critical island size. We offer a rough estimate of this spread to show that this can be the source of the dispersion of primary islands. Swartzentruber *et al.*²³ have carried out studies of these step fluctuations on Si(001) surfaces at 600 °C. We are interested in the amplitudes of fluctuations with wavelengths of the size of the depletion regions seen above ~ 1000 Å.⁶ Amplitudes are of the order of 100 Å for such fluctuations. We arrive at the same figure using the results of Bartelt *et al.*²⁴ Similar sized amplitudes are found in fluctuations of these wavelengths on steps of Ge/Si(001) in temperatures from 400 to 600 °C.^{13,22,25} These figures are also consistent with the fluctuations seen in InAs/GaAs systems at 520 °C.⁶ If we assume that the average depletion area has a radius of 500 Å,⁶ then a fluctuation in the radius of 100 Å would give island size distributions with total dispersions ($\Delta h/h + 2\Delta s/s$) of up to 40%, comparable to the experimental figures of Table I.

In contrast to the results of experiments discussed above,⁴⁻⁶ Polimeni *et al.* report a smooth 2D–3D transition for the growth of InAs on GaAs(001). However, their growth temperature at 420 °C is substantially lower than those (500–530 °C) used by the other groups. At these temperature differences, the diffusion constant D could differ by an order of magnitude, if we assume similar behavior for D observed for Si atoms on Si(001).^{16,17} Then the deposition rate of 0.1 ML/s used by Polimeni *et al.* is effectively much higher than those used by Leonard *et al.*⁴ and Moison *et al.*⁵ As we have already noted for the Ge/Si system, high deposition rates may mask the phenomenon of depletion and so give rise to a smooth 2D–3D transition.

We now consider the temperature-concentration phase curve, delineating smooth from rough growth, for the deposition of $\text{Si}_{1-x}\text{Ge}_x$ on Si(001) obtained by Bean *et al.*²⁶ As seen above, islands nucleate and grow two dimensionally until they reach a critical size s_c , when it becomes energetically favorable for the islands to grow three dimensionally. For the transition to occur, the temperature T must be high enough or flux rate R sufficiently low. Consider the following equation for the energy E of an island of side s and height h ,^{8,14}

$$E/V = \frac{2\Gamma}{s} - \frac{2c[1 - \exp(-\alpha h/s)]}{\alpha \ln[se^{1.5}/h \cot(\theta)]}, \quad (1)$$

where $\Gamma = \gamma_e \csc(\theta) - \gamma_s \cot(\theta)$, $\gamma_{s,e}$ are, respectively, energies of the surface and the facet, θ , the angle between the surface and facet, α , some constant, and $c \sim x^2$. Equation (1) shows that $s_c \sim c^{-1} \sim x^{-2}$. For an island of size s_c to form, an adatom must be able to diffuse this distance before capture. If τ is the lifetime of the adatom, then $s_c \sim \sqrt{D\tau}$, where $\tau \sim R^{-1}$ and $D = D_0 \exp(-E_d/kT)$. Then $s_c \sim \exp(-E_d/kT) \sqrt{\tau}$ or $x_c \sim \tau^{-1/4} \exp(E_d/4kT)$. In Fig. 3,

we replot Fig. 1 of Bean *et al.*²⁶ in the form of $1/T$ against $\ln(x)$. Except for the point at $x = 1$ (where a change in temperature from 550 to 527 °C would put it on the line) our results agree well with the experimental observations.

In conclusion, we find that in general, for strained heteroepitaxial growth of semiconductors, there exists an effective barrier for the 2D–3D transition. This is a robust result that should apply to a wide range of semiconductor systems. For highly mismatched systems, it is the underlying reason for the uniformity in the sizes of islands seen experimentally. It is consistent with other experimental results; the increase in island density with coverage with no corresponding increase in size; the phenomenon of depletion; the (initially unexpected) result that island density increases with reduced growth rate, or enhanced diffusion. Our results lead to a film morphology versus temperature and concentration phase diagram that is in qualitative agreement with experimental results. In a future publication,¹² we will show that it gives some insight into the nature of the rippling seen in low x systems.²⁷

Acknowledgment: This work is supported by the U.S. Office of Naval Research.

¹Y.-W. Mo, D. E. Savage, B. S. Swartzentruber, and M. G. Lagally, *Phys. Rev. Lett.* **65**, 1020 (1990).

²P. O. Hansson, M. Albrecht, H. P. Strunk, E. Bauser, and J. H. Werner, *Thin Solid Films* **216**, 199 (1992).

³M. Sopanen, H. Lipsanen, and J. Ahopelto, *Appl. Phys. Lett.* **67**, 3768 (1995).

⁴D. Leonard, M. Krishnamurthy, C. M. Reaves, S. P. Denbars, and P. M. Petroff, *Appl. Phys. Lett.* **63**, 3203 (1993); D. Leonard, K. Pond, and P. M. Petroff, *Phys. Rev. B* **50**, 11687 (1994).

⁵J. M. Moison, F. Houzay, F. Barthe, and L. Leprince, *Appl. Phys. Lett.* **64**, 196 (1994).

⁶J. M. Gerard, J. B. Genin, J. Lefebvre, J. M. Moison, N. Lebouche, and F. Barthe, *J. Cryst. Growth* **150**, 351 (1995); J. M. Gerard, O. Cabrol, J. Y. Marzin, N. Lebouche, and J. M. Moison, *Mater. Sci. Eng. B* **37**, 9 (1996).

⁷Bert Voigtlander and Andre Zinner, *Appl. Phys. Lett.* **63**, 3055 (1993); personal communication.

⁸K. E. Khor and S. Das Sarma, *Phys. Rev. B* **49**, 13657 (1994); *ibid.* **50**, 18382 (1994).

⁹D. J. Chadi, *Phys. Rev. Lett.* **59**, 1691 (1987).

¹⁰C. Priester and M. Lannoo, *Phys. Rev. Lett.* **75**, 93 (1995).

¹¹Atsushi Oshiyama, *Phys. Rev. Lett.* **74**, 130 (1995).

¹²K. E. Khor and S. Das Sarma (unpublished).

¹³F. Iwawaki, M. Tomitori, and O. Ishikawa, *Surf. Sci. Lett.* **253**, L411 (1991).

¹⁴J. Tersoff and R. M. Tromp, *Phys. Rev. Lett.* **70**, 2782 (1993).

¹⁵K. E. Khor, *Bull. Am. Phys. Soc.* **40**, 621 (1995).

¹⁶Y.-W. Mo, J. Kleiner, M. B. Webb, and M. G. Lagally, *Surf. Sci.* **268**, 275 (1992).

¹⁷J. A. Venables, *Surf. Sci.* **299/300**, 798 (1994).

¹⁸Y.-W. Mo and M. G. Lagally, *Mater. Sci. Eng. B* **14**, 311 (1992); J. Knall and J. B. Pethica, *Surf. Sci.* **265**, 156 (1992).

¹⁹D. Leonard, M. Krishnamurthy, S. Fafard, J. L. Merz, and P. M. Petroff, *J. Vac. Sci. Technol. B* **12**, 1063 (1994).

²⁰G. S. Solomon, J. A. Trezza, and J. S. Harris, Jr., *Appl. Phys. Lett.* **66**, 3161 (1995).

²¹F. Iwawaki, M. Tomitori, and O. Ishikawa, *Ultramicroscopy* **42-44**, 902 (1992).

²²U. Kohler, O. Jusko, B. Muller, M. Horn-von Hoegen, and M. Pook, *Ultramicroscopy* **42-44**, 832 (1992).

²³B. S. Swartzentruber, Y.-W. Mo, R. Kariotis, M. G. Lagally, and M. B.

- Webb, Phys. Rev. Lett. **65**, 1913 (1990).
- ²⁴N. C. Bartelt, R. M. Tromp, and Ellen D. Williams, Phys. Rev. Lett. **73**, 1656 (1994).
- ²⁵Fang Wu, Xun Chen, Zhenyu Zhang, and M. G. Lagally, Phys. Rev. Lett. **74**, 574 (1995).
- ²⁶J. C. Bean, T. T. Sheng, L. C. Feldman, A. T. Fiory, and R. T. Lynch, Appl. Phys. Lett. **44**, 102 (1984).
- ²⁷A. J. Pidduck, D. J. Robbins, A. G. Cullis, W. Y. Leong, and A. M. Pitt, Thin Solid Films **222**, 78 (1992); A. G. Cullis, D. J. Robbins, S. J. Barnett, and A. J. Pidduck, J. Vac. Sci. Technol. A **12**, 1924 (1994).

Morphological stability of strained-layer semiconductors

B. W. Wessels^{a)}

Department of Materials Science and Engineering, and the Materials Research Center,
Northwestern University, Evanston, Illinois 60208

(Received 13 January 1997; accepted 16 April 1997)

The morphological stability of strained-layer semiconductors is analyzed for the case where strain relaxation occurs by formation of isolated, coherent islands. In the model, partial strain relaxation in the underlying strained-layer is taken into account. Using a mean-field approximation, the interaction energy is taken to be linearly proportional to the strain energy in the film. The model predicts that the critical thickness for island formation depends inversely on the square of the misfit. The predicted dependence of critical thickness on the misfit is in quantitative agreement with recent experimental studies on island nucleation for heteroepitaxial III-V compounds. © 1997 American Vacuum Society. [S0734-211X(97)08304-2]

I. INTRODUCTION

Strained-layer III-V heterostructures are of considerable interest for a variety of electronic and optoelectronic device applications. Of particular interest are lower-dimensional structures such as quantum wells, wires, and dots. However, highly strained structures are difficult to prepare since strain leads to nonplanar growth.¹⁻⁵ The growth generally proceeds via a Stranski-Krastanow (SK) growth mode whereby after an initial planar growth, formation of islands occurs. These islands are often coherent as demonstrated by transmission electron microscopy (TEM).^{6,7} Formation of these islands provides a mechanism of local strain relief, thereby lowering the system energy. Metastable equilibrium models have been developed to describe the nucleation of discrete islands.⁸⁻¹⁰ In these models, the coherent islands are taken to nucleate on a wetting layer and are considered to be partially or fully relaxed. Island nucleation is taken as independent of the wetting layer thickness. This is in contrast to a number of experimental studies whereby it is observed that there is a finite critical thickness before island formation.^{3-5,11} The observed critical thickness depends on the composition of the wetting layer. Indeed, for growth of InGaAs on GaAs, it was shown that the critical thickness for island formation was inversely proportional to the square of the composition of the planar wetting layer.^{4,12} A basic question that remains is how does the strained epitaxial planar layer influence the subsequent nucleation of islands?¹³

In addition to TEM studies, recent finite element calculations have indicated that considerable elastic relaxation occurs beneath the coherent island within the strained layer and the substrate.¹⁴ Thus, calculations of the critical thickness should take into account the strained-layer contribution to overall strain relaxation. Towards this end, we developed a model whereby effects of relaxation of the underlying strained layer were taken into account using a mean-field approximation.¹² The model correctly describes the experimentally observed dependence of critical thickness on layer misfit. In this article, the formalism for the prediction of

island formation is further developed and an explicit expression for critical thickness is derived. Calculations for a prototype system are presented and compared to experiment.

II. THEORY OF COHERENT ISLAND NUCLEATION

For the nucleation of an isolated, coherent island from the vapor phase on top of a growing planar layer, the expression for the change in the energy is given by

$$\Delta G = \frac{2}{3}\pi r^3 \sigma_i + \pi r^2 \gamma_F + \Delta E_v, \quad (1)$$

where the first term is the volume-free energy of a relaxed, hemispherical island with respect to a flat, strained surface, and the second term is the energy of the additional free surface exposed when the island is formed. It has been assumed here, for convenience, that the surface energy per unit area γ_F of the strained film and the relaxed island is the same. The last term, ΔE_v , is an energy that results from the interaction between the lattice relaxed island and the strained layer through elastic deformation. Upon island nucleation, the underlying layer partially relaxes.¹⁴

The interaction energy can be obtained using a formalism similar to that previously developed by Matthews *et al.* for describing nucleation of incoherent islands on top of a strained planar wetting layer (see Fig. 1).¹⁵ Before island nucleation, the areal strain energy is given by

$$E = \mu \delta^2 h, \quad (2)$$

where μ is the modulus, δ is the misfit between the layer and the substrate, and h is the epitaxial layer thickness. After island nucleation, the underlying film is partially relaxed and the areal strain energy is then given by

$$E_r = \mu \epsilon^2 h, \quad (3)$$

where ϵ is the mean misfit strain in the strained layer that minimizes the total energy.^{15,16} It follows then that the net areal interaction energy is

$$\Delta E = (\epsilon^2 - \delta^2) \mu h = - \left(1 - \frac{\epsilon^2}{\delta^2} \right) \delta^2 \mu h \quad (4)$$

^{a)}Electronic mail: b-wessels@nwu.edu

or

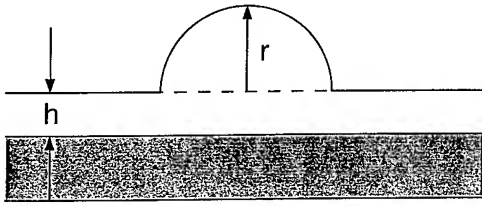


FIG. 1. Schematic of a hemispherical island nucleating on a strained layer.

$$\Delta E = -\lambda(\epsilon)E, \quad (5)$$

where $\lambda(\epsilon) = (1 - \epsilon^2/\delta^2)$. Note λ is always positive and less than 1, since ϵ^2 is less than δ^2 . Thus, the interaction energy is always negative. For the limiting case of complete layer relaxation, $\epsilon=0$ and λ is equal to 1. For the case of small relaxation, $\lambda = \lambda_0 + \lambda_1\epsilon + \lambda_2\epsilon^2$. For small strains $\lambda = \lambda_0$, a constant of the order of 1.

The volume interaction energy is, thus,

$$\Delta E_v = -\lambda\mu\delta^2 hA, \quad (6)$$

where the interaction volume Ah is taken as a cylinder of radius r and height h . Here, we are assuming that only the film underneath the island relaxes consistent with finite element analysis.¹⁴

Substituting for ΔE_v , the change in the energy is given by

$$\Delta G = \frac{2}{3}\pi r^3 \sigma_i + \pi r^2 \gamma_F - \lambda\mu\delta^2 h \pi r^2. \quad (7)$$

To determine the critical radius r_c , ΔG is differentiated with respect to r ,

$$r_c \sigma_i + \gamma_F - \lambda\mu\delta^2 h = 0, \quad (8)$$

and the critical radius of the island is then given by

$$r_c = -\left(\frac{\gamma_F - \lambda\mu\delta^2 h}{\sigma_i} \right). \quad (9)$$

In the limit as r_c approaches zero,

$$h = h_{c0} = \gamma_F / \lambda\mu\delta^2. \quad (10)$$

Equation (10) gives a thickness criterion for island formation for strained layers. This represents the critical thickness. The exact critical thickness would also depend on the island geometry. Combining terms in Eq. (7), the free energy is given by

$$\Delta G = \frac{2}{3}\pi r^3 \sigma_i + \pi r^2 \gamma_F^*, \quad (11)$$

where γ_F^* is an effective surface energy that could, in principle, be zero or positive depending on the strain in the layer, but $\gamma_F^* \leq \gamma_F$. The critical radius is then given by

$$r_c = -\gamma_F^* / \sigma_i. \quad (12)$$

From Eq. (8), the critical thickness for $r=r_c$ can be obtained by solving for h , and is given by

$$h_c = (\sigma_i r_c + \gamma_F^*) / \lambda\mu\delta^2 \quad (13)$$

and

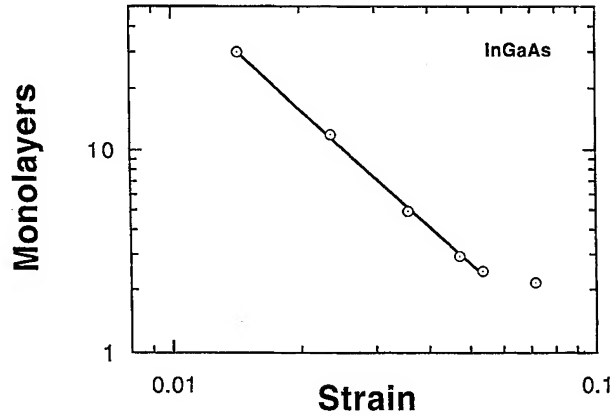


FIG. 2. Log of critical thickness versus log strain for epitaxial InGaAs on GaAs. The data are from Ref. 4. The solid line is calculated.

$$h_c = \gamma_{Fc}^* / \mu\delta^2, \quad (14)$$

where γ_{Fc}^* is the effective surface energy for a nucleus of $r=r_c$.

III. COMPARISON TO EXPERIMENT

To test the above expression, typical constants for a semiconductor were substituted into Eq. (14). The constants used are $\gamma_F = 1000$ ergs cm^{-2} , which is a typical energy for an unstrained semiconductor surface and for μ ,

$$\mu = \frac{2G(1+\nu)}{(1-\nu)}, \quad (15)$$

where G is the shear modulus [$G = 1 \times 10^{12}$ dynes/cm and ν is Poisson's ratio ($\nu = 0.33$)]. Upon substitution, a value of $\gamma/\mu = 2.7 \times 10^{-10}$ cm is obtained. For comparison, using the critical thickness data for the onset of island formation of Petroff and Den Baars (Ref. 4) for InGaAs on GaAs deposited by molecular beam epitaxy, a value of $\gamma/\mu = 1.8 \times 10^{-10}$ cm is calculated. The agreement is remarkably good even though the surface energy for an unstrained surface was used. Using this surface energy would lead to an overestimate. In terms of the unstrained surface energy, the critical thickness is approximately given by

$$h_c = \frac{\gamma_F}{\mu\delta^2} (1 - \lambda\mu\delta^2 h_c / \gamma_F). \quad (16)$$

Figure 2 shows the experimental data of Petroff and Den Baars for the onset of island formation and the model calculations. Excellent agreement between the data and the calculations is obtained. The constants used in the calculation were $G = 1 \times 10^{12}$ dynes cm^{-2} , $\nu = 0.33$ and $\gamma_F^* = 670$ ergs cm^{-2} . The difference between the unstrained and strained surface energies is 330 ergs cm^{-2} . This is appreciable and, thus, strain effects cannot be neglected in predicting the critical thickness for island formation.

Using these constants, the critical thickness for island formation is given by

$$h_c \delta^2 = 1.8 \times 10^{-10}, \quad (17)$$

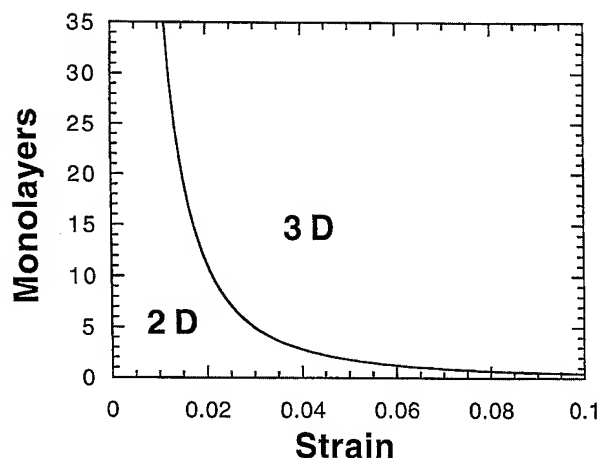


FIG. 3. Calculated critical thickness for roughening via coherent island formation as a function of strain.

where h_c is in cm. The critical thickness as a function of misfit strain is shown in Fig. 3. Note that while this expression is for the InGaAs/GaAs system, it should also approximate the behavior of other semiconductor systems since γ_F/μ is approximately constant for other semiconductors of the same class for the same surface orientation.¹⁷

IV. SUMMARY

In summary, the morphological stability of strained-layer semiconductor thin films was analyzed for the case where strain relaxation occurs by formation of isolated, coherent islands. In the model, partial strain relaxation in the underlying strained layer is taken into account using a mean-field approximation. The model predicts that the critical thickness for island formation depends inversely on the square of the misfit of the strained layer. The calculated critical thickness

is in excellent quantitative agreement with recent experiments on the onset of island formation in strained-layer III-V semiconductors.

ACKNOWLEDGMENTS

Useful discussions with J. Guyer, S. Theiring, M. Olvera, and P. Voorhees are gratefully acknowledged. This work is supported by the MRSEC program of the National Science Foundation at the Materials Research Center of Northwestern University under Award No. DMR-9632472.

- ¹R. P. Schneider and B. W. Wessels, *Appl. Phys. Lett.* **54**, 1142 (1989).
- ²R. P. Schneider, Jr. and B. W. Wessels, *Appl. Phys. Lett.* **57**, 1998 (1990).
- ³L. Q. Qian and B. W. Wessels, *Appl. Phys. Lett.* **63**, 628 (1993).
- ⁴P. M. Petroff and S. P. DenBaars, *Superlattices Microstruct.* **15**, 15 (1994).
- ⁵D. Leonard, M. Krishnamurthy, S. Fafard, J. L. Merz, and P. M. Petroff, *J. Vac. Sci. Technol. B* **12**, 1063 (1994).
- ⁶D. J. Eaglesham and M. Cerullo, *Phys. Rev. Lett.* **64**, 1943 (1990).
- ⁷S. P. Guha, A. Madhukar, and K. C. Rajkumar, *Appl. Phys. Lett.* **57**, 2110 (1990).
- ⁸J. Tersoff and R. M. Tromp, *Phys. Rev. Lett.* **70**, 2782 (1993).
- ⁹J. Tersoff and F. K. LeGoues, *Phys. Rev. Lett.* **72**, 3570 (1994).
- ¹⁰B. J. Spencer and J. Tersoff, *Mater. Res. Soc. Symp. Proc.* (to be published).
- ¹¹N. Grandjean, J. Massies, and F. Raymond, *Jpn. J. Appl. Phys. 2, Lett.* **33**, L1427 (1994).
- ¹²B. W. Wessels, *Mater. Res. Soc. Symp. Proc.* **440** (1997).
- ¹³N. Cabrera, *Surf. Sci.* **2**, 320 (1964).
- ¹⁴S. Christiansen, M. Albrecht, H. P. Strunk, and H. J. Maier, *Appl. Phys. Lett.* **64**, 3617 (1994).
- ¹⁵J. W. Matthews, D. C. Jackson, and A. Chambers, *Thin Solid Films* **26**, 129 (1975).
- ¹⁶The strain energy density has been calculated for the case of strained-layer epitaxy on finite size samples. S. Luryi and S. Suhir, *Appl. Phys. Lett.* **49**, 140 (1986).
- ¹⁷In the above treatment surface diffusion was sufficient so as not to limit island formation. When surface diffusion is rate limiting, the critical thickness for island formation is predicted to have a $(\gamma/\mu\delta^2)^2$ dependence. C. W. Snyder, J. F. Mansfield, and B. G. Orr, *Phys. Rev. B* **46**, 9551 (1992).

Second harmonic spectroscopy of Si(001) surfaces: Sensitivity to surface hydrogen and doping, and applications to kinetic measurements

Z. Xu,^{a)} X. F. Hu, and D. Lim

Department of Physics, The University of Texas at Austin, Austin, Texas 78712

J. G. Ekerdt

Department of Chemical Engineering, The University of Texas at Austin, Austin, Texas 78712

M. C. Downer

Department of Physics, The University of Texas at Austin, Austin, Texas 78712

(Received 12 January 1997; accepted 23 March 1997)

We report a spectroscopic second harmonic (SH) study of Si(001) surfaces under both ultrahigh vacuum and epitaxial growth conditions which includes *in situ* azimuthal rotation of the sample. The results show a strong influence of surface hydrogen termination and bulk doping concentration on the SH spectrum and azimuthal anisotropy. We propose a qualitative model in which a combination of bulk and surface electric-field induced SH polarizations, which augments the usual surface dipole and bulk quadrupole SH polarizations, are primarily responsible for the sensitivity to surface hydrogen and bulk doping. We also report an *in situ* real-time SH study of disilane adsorption onto and hydrogen desorption from Si(001) surfaces during epitaxial growth. The abilities to detect bulk doping concentration and monitor the growth rate in real time make SH spectroscopy promising as a growth sensor. © 1997 American Vacuum Society. [S0734-211X(97)08804-5]

I. INTRODUCTION

Noninvasive optical probes are becoming increasingly attractive both in the fundamental study of solid surfaces and in applications such as real-time monitoring and control of semiconductor epitaxial growth.¹ Optical second harmonic (SH) spectroscopy provides a surface-specific and surface-sensitive probe for centrosymmetric semiconductors such as silicon because the bulk second order susceptibility vanishes in the electric dipole approximation.² Indeed SH probing of surface desorption and surface structural phase transitions has been reported recently.³⁻⁵ A very recent study showed the sensitivity of a SH probe for measuring the very low sticking probability for dissociative adsorption of H₂ on Si(001) 2×1 surfaces.⁶ However, most of those studies were performed at a fixed wavelength with fixed sample orientation. In such a restricted parameter space, the physical origin of the surface SH generation is difficult to establish. This lack of understanding also inhibits further applications.

In this article, we report an *in situ* spectroscopic SH study of bare and H-terminated Si(001) surfaces using a Ti:sapphire femtosecond laser system that offers large wavelength tunability, together with high peak power necessary for strong nonlinear optical signals without significant sample heating. The article is presented in two parts: (1) A study of the phenomenology and physical origin of SH generation at bare, H-terminated, doped and undoped Si(001) surfaces. In this study, we exploit the source laser tunability and the ability to rotate the sample azimuthally in our ultrahigh vacuum (UHV) chamber, and thus to measure the strong azimuthal anisotropy, which is the hallmark of nonlinear optical probes.⁷ The observed sensitivity of the SH spectrum and

azimuthal anisotropy to surface H and bulk doping is explained by a combination of bulk and surface electric-field induced SH (EFISH) contributions, which supplement the usual surface dipole and bulk quadrupole contributions included by previous investigators. The sensitivity of the SH signal to bulk doping concentration has not been reported before in Si(001). (2) A real-time SH study of disilane adsorption and hydrogen desorption kinetics during silicon epitaxial growth on Si(001). This study demonstrates the ability of SH generation to monitor the growth rate in real time.

II. EXPERIMENTAL PROCEDURE

The experiments were performed in an UHV deposition chamber pumped by a 450 ℓ/s turbomolecular pump to a base pressure of 5×10^{-10} Torr, consisting mostly of molecular hydrogen, for which no special provision for pumping was made because of the weak dissociative adsorption of H₂ due to its low sticking coefficient ($< 10^{-6}$) on Si(001).^{8,9} Partial pressure of water and other gases was $< 5 \times 10^{-11}$ Torr, as indicated by residual gas analysis. The deposition chamber was equipped for reflection high energy electron diffraction (RHEED), and was adjoined via load lock to a multitechnique surface analysis chamber (base pressure 8×10^{-11} Torr), which was equipped with a quadrupole mass spectrometer (QMS) used to calibrate fractional hydrogen coverage θ_H by temperature-programmed desorption (TPD) following procedures discussed previously.¹⁰ Experiments were performed on two types of Si(001) samples, both with negligible miscut ($< 0.5^\circ$): (1) an undoped epitaxial silicon layer of approximately 0.2 μm thickness grown on a Si(001) substrate by UHV-chemical vapor deposition using disilane (Airco, electronic grade mixture consisting of 4%-Si₂H₆/balance He) with a gas phase pressure of 5

^{a)}Electronic mail: zxu@physics.utexas.edu

$\times 10^{-5}$ Torr at 650 °C for approximately 15 min. This film thickness exceeds the SH escape depth; (2) heavily *p*-type doped Si(001) (boron doped, 0.01–0.02 Ω cm resistivity). The RHEED pattern for the clean surfaces of all samples showed a sharp 2×1 reconstruction feature. Hydrogen was deposited by exposing the bare surface to atomic hydrogen produced by cracking molecular hydrogen on a 2000 K tungsten filament. The surface was dosed at 150 °C to suppress surface roughening caused by etching, which occurs in lower temperature dosing.¹¹ When working on unsaturated, hydrogen-terminated surfaces, we flashed to 1050 K and dosed hydrogen within approximately 15 min to avoid surface contamination. The surface was monitored with RHEED following each exposure and SH experiment and sharp 2×1 reconstruction features were maintained during the entire experimental sequence.

We used laser pulses of 120 fs duration, 5–10 nm spectral bandwidth [full width at half-maximum (FWHM)], 500 mW average power, 76 MHz repetition rate, and 710–800 nm tuning range from an unamplified Ti sapphire laser for SH spectroscopy. The laser beam was focused at $f/100$ through a fused-silica viewport onto the silicon crystal *in vacuo* at a 55° angle of incidence with *p* polarization. The signal collection optics imaged *p*-polarized SH light from the irradiated sample spot through another optical viewport onto the center of the cathode of a photomultiplier tube equipped with gated photon counting electronics. The sample was rotated through a 240° azimuthal angle by a mechanical feed through with less than $\pm 3^\circ$ residual wobble of the surface normal. The imaging collection optics kept the SH light fixed within 1 mm² area on the photocathode despite this residual wobble, thus eliminating the major source of rotational artifacts in the data. We confirmed that the SH signal was insensitive to the slight variations in angle of incidence, which also resulted from this wobble. Part of the beam was focused onto a quartz crystal to generate a reference SH signal to normalize the effects of laser output variation. Other details of the apparatus are described elsewhere.¹⁰

We performed two types of experiments. In the first part (mechanism study), we performed SH studies with azimuthal sample rotation on Si(001) surfaces with various hydrogen coverage and two doping concentrations. In the second part (application study), we studied hydrogen desorption kinetics by performing real-time, fixed wavelength measurements after disilane dosing at various sample temperatures, and studied disilane adsorption kinetics by monitoring the steady-state SH intensity at various disilane gas pressures.

III. SH MECHANISM STUDY

A. Theoretical background

The SH polarization at Si(001) interfacial regions for a *p*-in/*p*-out configuration can be phenomenologically expressed as¹²

$$P(2\omega) = P^{\text{isotropic}} + P^{\text{BQ}} \cos 4\varphi, \quad (1)$$

where φ is the azimuthal angle between the incident plane and the $\langle 110 \rangle$ crystalline direction, $P^{\text{isotropic}}$ represents an azi-

mutually isotropic SH dipole polarization, and P^{BQ} is the fourfold anisotropic bulk quadrupole polarization that becomes important because the electric dipole polarization vanishes in the centrosymmetric silicon bulk. $P^{\text{isotropic}}$ has usually been attributed almost entirely to a surface polarization P^S , caused by inversion symmetry breaking at the surface. Early pioneering SH spectroscopy by Daum *et al.* suggested that dimer-induced vertical strain in the subsurface region of Si(001) 2×1 was in fact the dominant contribution to P^S , since saturation hydrogen coverage (1.5 ML), which relaxes this strain, was observed to quench the SH signal.¹³ Based on this model, the strongest SH spectroscopic changes on Si(001) surfaces should be expected for coverages exceeding one monolayer, because the dimer-induced subsurface vertical strain is only slightly affected by submonohydride termination but relaxes almost completely with dihydride termination.¹⁴ This expectation however was challenged by a recent study, in which much stronger SH spectroscopic changes were observed with coverages below monohydride termination than above monohydride termination,¹⁰ suggesting that a different mechanism was at work.

Some recent SH studies revealed the importance of a DC electric-field induced SH (EFISH) polarization $P^E \sim \chi^{(3)} E_{\text{dc}} E^2(\omega)$ at Si interfaces, which had been neglected in most previous studies.¹⁵ Such dc fields, which break inversion symmetry, are presented in the space charge region (SCR) that arises when interface states, e.g., on bare silicon surfaces, pin the surface Fermi energy close to midgap.¹⁶ This dc electric field penetrates to a distance below the surface determined by the bulk doping concentration.¹⁷ As a result, this bulk EFISH contribution P^{BE} exhibits essentially bulk spectroscopic properties. For example, its spectral peaks correspond to bulk critical point resonances (such as E_1 resonance at 3.4 eV).^{15,18} Because P^{BE} , like P^S , is isotropic with respect to azimuthal sample rotation, we treat P^{BE} formally as a component of $P^{\text{isotropic}}$ in Eq. (1). Strong dc electric fields are also present within the top 2 to 3 atomic monolayers (ML) of the reconstructed Si surface because of local valence electron redistribution into buckled surface dimers.¹⁷ The resulting surface EFISH contribution P^{SE} is also rotationally isotropic and thus will also be treated as an effective component of $P^{\text{isotropic}}$. Thus we can write

$$P^{\text{isotropic}} = P^S + P^{\text{BE}} + P^{\text{SE}}. \quad (2)$$

Separation of the various SH sources described above requires measurements that separately modify the individual contributions. In our study, we quench the bulk and surface EFISH contributions simultaneously by hydrogen termination, then identify the role of bulk EFISH by varying the bulk doping concentration. Monohydride termination on Si surfaces quenches the bulk EFISH contribution by removing dangling bond states close to midgap. Monohydride termination quenches surface EFISH by redistributing valence electrons from dimers into the bulk.¹⁷ Thus, monohydride termination is expected to significantly modify both EFISH contributions while leaving subsurface vertical strain nearly

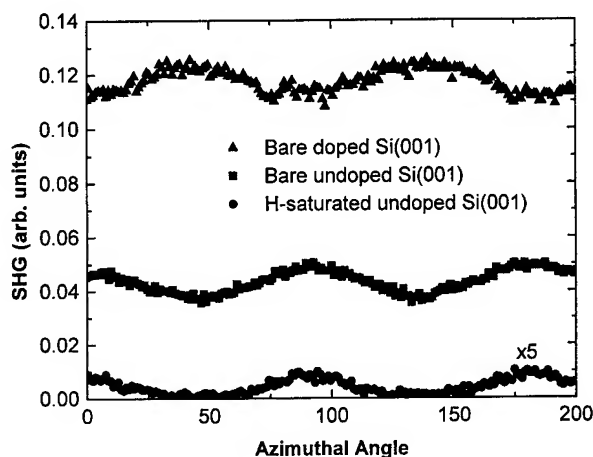


FIG. 1. Representative p -in/ p -out SH signals from Si(001) surfaces with incident wavelength of 740 nm showing variation in fourfold anisotropic and isotropic components with doping and H termination: clean p -type doped Si(001)- 2×1 ($p=5\times 10^{18}\text{ cm}^{-3}$) (triangles); clean epitaxial undoped Si(001)- 2×1 (squares); and hydrogen saturated ($\theta_H=1.5\text{ ML}$) undoped epitaxial Si(001) surface (circles). The signal from hydrogen saturated undoped sample is multiplied by 5. Note 45° azimuthal phase shift and three-fold amplitude ratio between doped and undoped clean surfaces.

unaffected. Doping, on the other hand, selectively modifies the bulk EFISH contribution by modifying both the surface band bending and SCR thickness.

The SH intensity generated by the SH polarization $P(2\omega)$ represented by Eq. (1) can be expressed phenomenologically as

$$I_{pp} \propto |P(2\omega)|^2 = a_0 + a_4 \cos(4\varphi) + a_8 \cos(8\varphi), \quad (3)$$

where a_0 (isotropic), a_4 (fourfold anisotropic), a_8 (eightfold anisotropic) are azimuthal Fourier coefficients, which can be related to the polarizations in Eq. (1) as follows:

$$\begin{aligned} a_0 &= |P^{\text{isotropic}}|^2 + (1/2)|P^{\text{BQ}}|^2, \\ a_4 &= 2 \text{Re}[P^{\text{isotropic}} \cdot P^{\text{BQ}*}], \\ a_8 &= (1/2)|P^{\text{BQ}}|^2. \end{aligned} \quad (4)$$

For Si(001) at wavelengths used in our study, P^{BQ} is always much smaller than $P^{\text{isotropic}}$. In fact, a_8 is below our noise level, and thus much smaller than the strong isotropic contribution a_0 . As a result of this inequality, the physical interpretation of a_0 and a_4 is simplified. Specifically, a_0 mainly reflects the property of $P^{\text{isotropic}}$, while a_4 depends on the magnitude of both $P^{\text{isotropic}}$ and P^{BQ} , as well as on their relative phase. Thus, the SH polarizations $P^{\text{isotropic}}$ and P^{BQ} at the Si(001) surface are straightforwardly related to a_0 and a_4 , which in turn, are derived directly from analysis of azimuthal scan data.

B. Experimental results

Figure 1 shows SH signals $I_{pp}(2\omega, \varphi)$ generated with an incident wavelength of 740 nm from (a) clean p -type doped, (b) clean undoped epitaxial, and (c) hydrogen saturated ($\sim 1.5\text{ ML}$) (undoped) samples as a function of azimuthal angle φ between the incident plane and the $\langle 110 \rangle$ crystalline

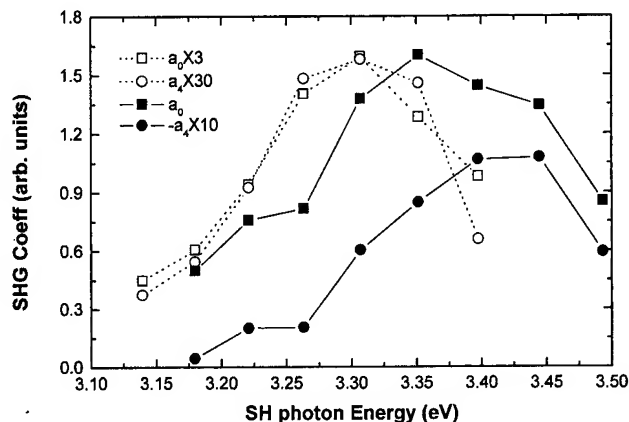


FIG. 2. Spectra of the isotropic a_0 (squares) and fourfold anisotropic a_4 (circles) SH components for the bare undoped (open symbols, dotted curves) and heavily p -type doped (filled symbols, solid curves) Si(001)- 2×1 surfaces. Note larger magnitude and bluer spectral peak for the doped sample.

direction. A fourfold oscillation (over 360°) superimposed on an approximately tenfold stronger isotropic background is evident in all cases. An eightfold anisotropic SH contribution also exists at Si(001), but was too weak to be observed in the present work. The data in Fig. 1 shows that the amplitude and azimuthal phase of the fourfold anisotropic and isotropic contributions depend on doping and θ_H . In particular, the signals from the clean p -type doped sample are several times stronger, and are shifted by 45° in azimuthal phase from the signals from the clean undoped sample. On the other hand, the signals from the hydrogen saturated p -type doped sample (not shown) have the same azimuthal phase and approximately the same amplitude as that from undoped samples.

To quantify the trends shown in Fig. 1, each azimuthal scan $I_{pp}(2\omega, \varphi)$ was fit to Eq. (3). For each surface, a_0 and a_4 were extracted for nine to eleven different frequencies near the two-photon E_1 resonance. The spectral dependence of a_0 and a_4 for the clean p -type doped and the clean undoped epitaxial surface is plotted in Fig. 2. For the undoped sample, a_0 and a_4 both show a single broad peak at $2h\nu = 3.3\text{ eV}$, which is consistent with the resonant feature at both Si/SiO₂ interfaces and bare Si surfaces observed by several authors.^{10,13,15} For the heavily doped sample, on the other hand, the peak maxima of a_0 and a_4 blueshifts to near $2h\nu = 3.4\text{ eV}$, and becomes 2 to 3 times stronger in magnitude. Note that a_4 has an opposite sign for the doped and undoped samples, consistent with the observed 45° azimuthal phase difference.

The spectral dependence of a_0 and a_4 for the undoped sample for five different H coverages is plotted in Fig. 3. In the top panel, the data for the bare surface is repeated for comparison. As θ_H increases from 0 to 1.5 ML, we observe three important trends: (1) the magnitudes of both a_0 and a_4 are gradually and continuously quenched by a factor of approximately seven, but retain their approximately tenfold ratio of magnitudes; (2) the spectral peak of a_0 red shifts continuously from $2h\nu = 3.3$ to 3.2 eV ; (3) the spectral peak of a_4 remains nearly fixed at $2h\nu = 3.3\text{ eV}$. The quenching

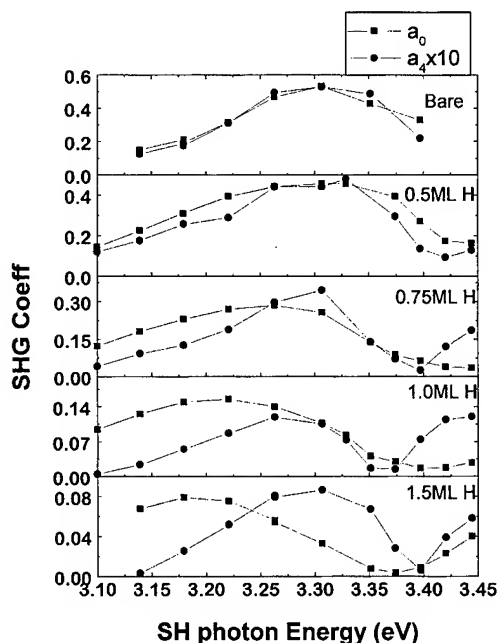


FIG. 3. Spectra of the isotropic a_0 (squares) and fourfold anisotropic a_4 (circles) SH contributions from an epitaxial undoped Si(001) surface as H coverage increases from 0 (top panel) to 1.5 ML (bottom panel). Note gradual red shift of a_0 spectrum, lack of shift of a_4 spectrum, and sevenfold quenching of amplitudes of both spectra.

and red shifting of the strong a_0 feature is consistent with the θ_H dependence of the fixed-azimuth spectra reported by Dadap *et al.*¹⁰ However, the unshifting a_4 spectrum is a previously unobserved feature.

Increasing θ_H on the doped sample quenches and red shifts the a_0 peak, quenches the a_4 peak, and reverses the sign of a_4 , producing SH spectra for saturation hydrogen coverage similar to those of the undoped sample as shown in Fig. 1 bottom.

C. Discussion

Increasing p -type doping not only shifts the bulk Fermi level towards the valence band but also decreases the thickness of the SCR, thus significantly enhancing the dc electric field due to surface band bending. This bulk dc electric field can be estimated using a Schottky barrier model: the SCR thickness is $\sim d_{\text{SCR}} = L_D \sqrt{2|\nu_s|}$, where L_D is the Debye length, which is $> 0.6 \mu\text{m}$ for the undoped sample and $\sim 6 \text{ nm}$ for the doped sample, and $\nu_s = eV_s/kT$ is the normalized surface band bending, which is much larger than unity in our case. For surface Fermi level pinning at $\sim 0.32 \text{ eV}$ above the valence band maximum, the band bending is $\sim -0.2 \text{ eV}$ for the undoped sample and $\sim 0.2 \text{ eV}$ for the doped sample. As a result, the dc electric field strength is $< 0.6 \text{ kV/cm}$ in the undoped sample and $\sim 60 \text{ kV/cm}$ in the heavily doped sample. This suggests that the bulk EFISH contribution may be negligible in the undoped sample, but can be very important in the heavily doped sample. Based on this argument, we can attribute the much stronger SH intensity of the doped sample shown in Fig. 2 to the doping en-

hanced bulk EFISH contribution. Equally importantly, the peak shifting towards 3.4 eV is also consistent with the expectation that bulk $\chi^{(3)}$ should show a strong bulk E_1 resonance feature, which is resonant at 3.4 eV .

On the other hand, bulk EFISH in the undoped sample is far too small to explain the signal level observed (only 2 to 3 times smaller than that from the heavily doped sample). Further evidence against the importance of the bulk EFISH contribution in the undoped sample comes from the gradual quenching of the SH signal as hydrogen coverage increases. The surface Fermi level pinning depends on the density of midgap surface states in a steplike function.¹⁷ As a result, the SH signal from the SCR should show a sudden quenching as hydrogen coverage increases, which clearly contradicts the gradual quenching shown in Fig. 3. The quenching and red shifts of the SH signal by monohydride termination also rules out subsurface vertical strain as the origin of the SH generation because the subsurface strain is nearly unchanged with monohydride termination, as discussed earlier.

Recent calculations show strong charge transfer resulting from dimerization on the bare Si(001) surface.¹⁹ This charge transfer can generate strong near-surface local electric fields,¹⁷ as evidenced by the observed silicon core-level shift on bare silicon surfaces.¹⁷ Assuming the core-level shift is caused completely by the near-surface local electric field (i.e., neglecting the contribution of near surface strain), the near surface field can be estimated as $\sim 0.5 \times 10^7 \text{ V/cm}$,¹⁷ which is consistent with the calculated valence electron redistribution.¹⁹ Although this electric field penetrates only ~ 2 to 3 monolayers below the surface, it is $\sim 10^4$ times larger than the bulk dc electric field from band bending in undoped sample. Thus it can generate a SH signal comparable to the bulk EFISH contribution in the heavily doped sample, assuming surface EFISH is governed by a third order susceptibility similar to that in the bulk. Most importantly, this near-surface local electric field can be quenched by monohydride termination in a gradual way. In addition, it is opposite in direction from the bulk band bending field in the heavily doped sample. This explains the opposite sign of a_4 we observed in doped and undoped samples, if we assume that bulk EFISH dominates in the doped sample while surface EFISH dominates in the undoped sample. Finally, hydrogen termination quenches both bulk and surface EFISH. As a result, SH spectra on hydrogen saturated doped and undoped samples are expected to be similar, as observed in our experimental results. Thus all major qualitative SH spectroscopic dependences on H coverage and doping are explained by the combined action of bulk and surface EFISH.

A more quantitative model of the SH spectra in Figs. 2 and 3 can be formulated by taking into account the bulk E_1 resonance at 3.4 eV and a red-shifted surface resonance. Surface SH is complicated by the presence of surface states. A recent spectroscopic study reveals two resonant peaks at ~ 3.2 and $\sim 3.4 \text{ eV}$, respectively.²⁰ The 3.2 eV resonance may be caused by a surface state enhanced E_0 resonant feature (i.e., the near doubly resonant transition: valence band \rightarrow surface states \rightarrow conduction band). While the exact

physical origin of this red-shifted resonant feature is still not clear, our observation of the peak shifting towards 3.2 eV as hydrogen coverage increases is consistent with their observation of this resonant feature. The a_0 peak shifting can be explained as the quenching of surface EFISH (~ 3.4 eV resonance) by hydrogen termination to reveal the 3.2 eV resonant feature by hydrogen termination. Numerical calculations²¹ based on the existence of two resonant features (3.2 eV and surface EFISH at ~ 3.4 eV) not only confirm this picture, but also reproduce the unshifted a_4 . The latter point can be qualitatively understood as the result of phase differences between different polarizations. As shown in Eq. (4), only the polarizations which are in phase with the bulk quadrupole polarization show up in a_4 . Further details of this model will be presented in a planned forthcoming article.²¹

IV. KINETICS STUDY

A. Theoretical background

The sensitivity of SH generation to surface hydrogen coverage makes SH spectroscopy a promising *in situ* monitor of chemical vapor deposition (CVD) epitaxial growth processes, in which hydrogen plays a crucial kinetic role. Success has been achieved by various groups in employing SHG to probe hydrogen desorption kinetics on silicon surfaces.^{4,6} However, because of the difficulty of determining hydrogen coverage at growth conditions, which is usually at a high temperature, real-time monitoring of growth rate has not been achieved. In our study, we demonstrate the feasibility of such a measurement when appropriate adsorption and desorption kinetics are independently determined.

For CVD growth employing disilane as the precursor, the surface reaction is quite complicated. However, a satisfactory phenomenological model can be based on the kinetic equation for competing surface hydrogen desorption and adsorption. Assuming first order hydrogen desorption and second order dissociative adsorption, which are consistent with previous studies of the growth mechanism,^{22,23} we have the governing reaction equation for the hydrogen coverage

$$d\theta_H/dt = (JS)(1 - \theta_H)^2 - k_d\theta_H, \quad (5)$$

where J is the incident disilane flux which is proportional to the disilane gas partial pressure P , S is the phenomenological reactive sticking coefficient, k_d is the desorption rate. Stoichiometric coefficients for the individual surface reactions that comprise dissociative adsorption and surface hydride formation are included with S . The first term thus represents adsorption and the second term represents desorption, which are balanced during steady-state growth. As a result, the growth rate is given by

$$GR = \eta k_d \theta_H (\text{ML/s}) = 1.4 \eta k_d \theta_H (\text{\AA/s}), \quad (6)$$

where η is a constant depending on the surface reactions. Assuming that all the hydrogen atoms in a disilane molecule stick to silicon sites during adsorption, we have three hydrogen atoms adsorbed for each adsorbed silicon atom, so that $\eta = 1/3$.

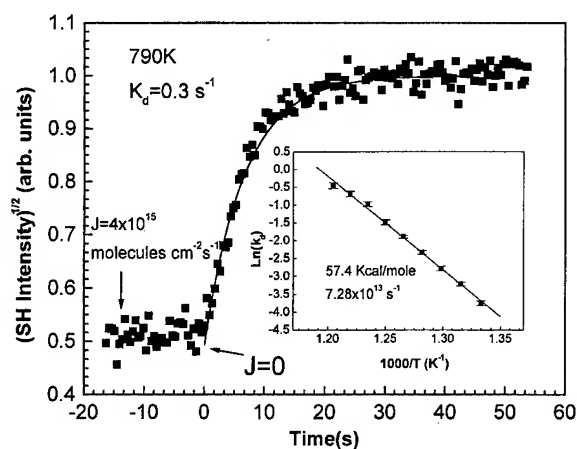


FIG. 4. The real-time SH response during isothermal hydrogen desorption at 790 K. The surface is initially dosed with 4% $\text{Si}_2\text{H}_6/\text{He}$ at flux $J = 4 \times 10^{15} \text{ mol cm}^{-2} \text{ s}^{-1}$. At $t = 0$, the flux is turned off within 0.5 s, which is the response time of the ion gauge. The desorption is fitted to an exponential growth function with $k_d = 0.3 \text{ s}^{-1}$. Desorption rates k_d at different temperatures are plotted in Arrhenius fashion in the inset. Fitting shows an activation energy (57.4 kcal/mol) and pre-exponential factor ($7 \times 10^{13} \text{ s}^{-1}$).

The key issue in employing SHG to monitor surface hydrogen coverage during growth is to establish the relation between the SH intensity $I(2\omega)$ and θ_H . Our room temperature studies on Si(001) surfaces with calibrated θ_H shows that $\sqrt{I(2\omega)} \propto \theta_H$ near the E_1 resonance. In our study, we assume that this relationship is true also at a high temperature, at which E_1 resonance red shifts to coincide with the incident wavelength of $\sim 800 \text{ nm}$ as evidenced by the previous temperature dependence study.¹⁰ However, the quantitative relation cannot be established as straightforwardly as at room temperature because of the difficulties in calibrating the hydrogen coverage at growth temperatures. In order to circumvent this difficulty, we measured SH intensity for a fixed wavelength (800 nm) at steady-state hydrogen coverage as a function of steady-state gas pressure P . From Eq. (5), the function can be expressed as follows:

$$P = \eta_s \theta_H / (1 - \theta_H)^2, \quad (7)$$

where η_s is a coefficient related to S , and $\sqrt{I(2\omega)} = \xi \theta_H + \xi_0$ with $\xi_0 = \sqrt{I(2\omega)}$ on bare surface. We can obtain ξ simply by fitting Eq. (7) to the experimental data and monitoring the hydrogen coverage with the SH intensity. In order to obtain the growth rate, we measured k_d by an isothermal study, in which desorption after hydrogen dosing is monitored. With flux J equal to zero, the solution of Eq. (5) is simply an exponential decay with decay time $1/k_d$. The temperature dependence of k_d can be described by a pre-exponential factor and an activation energy in the Arrhenius form $k_d = f_0 e^{-E_D/kT}$. For a high temperature (above 825 K), desorption cannot be measured directly because the desorption time constant is shorter than our time resolution (0.5 s). Thus we extrapolate the Arrhenius formula to those high growth temperatures. The details of the isothermal measurement technique have been discussed elsewhere.²⁴

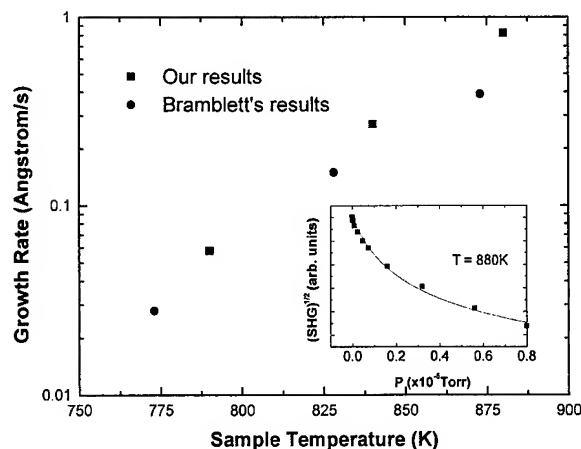


FIG. 5. The Arrhenius plot of calculated epitaxial growth rate (squares) for dosing with 4% $\text{Si}_2\text{H}_6/\text{He}$ at flux $J = 5 \times 10^{15} \text{ mol cm}^{-2} \text{ s}^{-1}$. The results are consistent with the results obtained by Bramblett *et al.* The SH intensity during steady-state measurements at sample temperature $T = 880 \text{ K}$ under various P ranged from low pressure up to $1 \times 10^{-5} \text{ Torr}$ is shown in the inset. The fitting of the steady state measurement result shown as the solid line gives the important ξ parameter that relates the SH intensity to the steady-state θ_H .

B. Experimental results and discussion

We performed isothermal measurements at temperatures ranging from 740 to 810 K. A typical result measured at $T = 790 \text{ K}$ is shown in Fig. 4. The measured k_d fits very well to an exponential function of $1/T$, as shown in the inset of Fig. 4. The pre-exponential factor of $7 \times 10^{13}/\text{s}$ and an activation energy of 57.3 kcal/mol agree well with values obtained by reflectance differential spectroscopy (RDS)²⁵ and with values determined by temperature programmed desorption.²⁶ We then performed steady-state measurements at three different temperatures, where the gas phase pressure was adjusted from low pressure up to $\sim 1 \times 10^{-5} \text{ Torr}$. Typical results measured at $T = 880 \text{ K}$ are shown in the inset of Fig. 5. By fitting the results to Eq. (7), we obtain the ξ coefficient which is critical in relating the SH intensity to the surface hydrogen coverage. With k_d and ξ obtained above, we calculated the growth rate at a flux of $5 \times 10^{15} \text{ mol cm}^{-2} \text{ s}^{-1}$ based on the SH intensity during steady-state growth. The results obtained at several temperatures are shown in Fig. 5. Our measurements are consistent with the results obtained by Bramblett *et al.*,²³ which are obtained at slightly higher flux. Most importantly, the SH signal can be obtained on a short time scale ($\sim 0.2 \text{ s}$) without interrupting the growth, which is essential for applications such as real-time active feed-back control of CVD growth. For monitoring changes (rather than absolute magnitudes) of growth rate in real time, absolute calibration of ξ is not critical. As a result, the monitoring can be even more flexible and easily applied. More complete measurements are in progress and will be reported elsewhere.²⁷

V. CONCLUSION

In summary, we reported our spectroscopic SH study of Si(001) in two parts: First, we studied the phenomenology and physical origin of SH generation from bare,

H-terminated, doped and undoped Si(001) surfaces. Our results show the sensitivity of SH spectroscopy on silicon surfaces to surface hydrogen coverage and bulk doping. We propose a simple qualitative model that attributes the SH sensitivity to surface H and bulk doping to a combination of bulk and surface EFISH contributions, which prominently augments the conventional surface dipole and bulk quadrupole contributions. Second, we demonstrate the ability to apply SH spectroscopy to *in situ*, real-time monitoring of growth rate, which is an important parameter for growth control. The combination of the abilities to detect bulk doping concentration and monitor the growth rate in real time makes SH spectroscopy very promising as a growth sensor.

ACKNOWLEDGMENTS

This research was supported by AFOSR/MURI Contract No. F49620-95-1-0475, AFOSR Contract No. F49620-95-C-0045, the NSF Science and Technology Center Program (Grant No. CHE-890210), the Robert Welch Foundation (Grant No. F-1038), and the Texas Advanced Technology Program (Grant No. ATPD-354).

¹D. E. Aspnes, *Surf. Sci.* **307–309**, 1017 (1994).

²Y. R. Shen, *Nature (London)* **337**, 519 (1989).

³K. Pedersen and P. Morgan, *Phys. Rev. B* **52**, R2277 (1995).

⁴T. F. Heinz, M. M. T. Loy, and W. A. Thompson, *Phys. Rev. Lett.* **54**, 63 (1985).

⁵H. B. Jiang, Y. H. Liu, X. Z. Lu, W. C. Wang, J. B. Zhang, and Z. M. Zhang, *Phys. Rev. B* **50**, 14 621 (1994).

⁶P. Bratu and U. Höfer, *Phys. Rev. Lett.* **74**, 1625 (1995); P. Bratu, K. L. Kompe, and U. Höfer, *Chem. Phys. Lett.* **1–7**, 251 (1996).

⁷B. Koopmans, F. van der Woude, and G. A. Sawatzky, *Phys. Rev. B* **46**, 12 780 (1992).

⁸G. Schulze and M. Henzler, *Surf. Sci.* **124**, 336 (1983).

⁹P. Bratu, W. Brinig, A. Grob, M. Hartmann, U. Höfer, P. Kratzer, and R. Russ, *Phys. Rev. B* **54**, 5978 (1996).

¹⁰J. I. Dadap, N. M. Russell, X. F. Hu, O. A. Aktsipetrov, J. G. Ekerdt, and M. C. Downer, *Phys. Rev. B* (in press).

¹¹S. M. Gates, R. R. Kurtz, and C. N. Greenlief, *Surf. Sci.* **207**, 364 (1989).

¹²J. E. Sipe, D. J. Moss, and H. M. van Driel, *Phys. Rev. B* **35**, 1129 (1987).

¹³W. Däum, H.-J. Krause, U. Reichel, and H. Ibach, *Phys. Rev. Lett.* **71**, 1234 (1993).

¹⁴Z. Jing and J. L. Whitten, *Phys. Rev. B* **46**, 9544 (1992).

¹⁵J. I. Dadap, X. F. Hu, M. H. Anderson, M. C. Downer, J. K. Lowell, and O. A. Aktsipetrov, *Phys. Rev. B* **53**, R7607 (1996).

¹⁶W. Mönch, P. Koke, and S. Kruger, *J. Vac. Sci. Technol.* **19**, 313 (1981).

¹⁷W. Mönch, *Semiconductor Surfaces and Interfaces* (Springer, Berlin, 1995).

¹⁸P. Godefroy, W. de Jong, C. W. van Hasselt, M. A. C. Deviller, and Th. Rasing, *Appl. Phys. Lett.* **68**, 1981 (1996).

¹⁹B. I. Craig and P. V. Smith, *Surf. Sci. Lett.* **226**, L55 (1989).

²⁰T. Yasuda, L. Mantese, U. Rossow, and D. E. Aspnes, *Phys. Rev. Lett.* **74**, 3431 (1995).

²¹Z. Xu, X. F. Hu, J. G. Ekerdt, and M. C. Downer, *Phys. Rev. B* (to be published).

²²U. Höfer, L. Li, and T. F. Heinz, *Phys. Rev. B* **45**, 9485 (1992).

²³T. R. Bramblett, Q. Lu, T. Karasawa, S. K. Jo, and J. E. Greene, *J. Appl. Phys.* **76**, 1884 (1994).

²⁴J. I. Dadap, X. F. Hu, N. M. Russell, J. G. Ekerdt, J. K. Lowell, and M. C. Downer, *IEEE J. Sel. Topics Quantum. Electron.* **1**, 1145 (1995).

²⁵J. W. Sharp and Gyula Eres, *Surf. Sci.* **320**, 169 (1994).

²⁶N. H. Russell and J. G. Ekerdt, *Surf. Sci.* (in press), and references therein.

²⁷X. F. Hu, Z. Xu, P. S. Parkinson, D. Lim, B. Gong, G. Hess, J. G. Ekerdt, and M. C. Downer (unpublished).

On the mechanism of the hydrogen-induced exfoliation of silicon

M. K. Weldon, V. E. Marsico, Y. J. Chabal, A. Agarwal,^{a)} D. J. Eaglesham, J. Sapjeta, W. L. Brown, D. C. Jacobson, Y. Caudano, S. B. Christman, and E. E. Chaban
Bell Laboratories, Lucent Technologies, 600 Mountain Avenue, Murray Hill, New Jersey 07974

(Received 12 January 1997; accepted 16 April 1997)

We have investigated the fundamental mechanism underlying the hydrogen-induced exfoliation of silicon, using a combination of spectroscopic and microscopic techniques. We have studied the evolution of the internal defect structure as a function of implanted hydrogen concentration and annealing temperature and found that the mechanism consists of a number of essential components in which hydrogen plays a key role. Specifically, we show that the chemical action of hydrogen leads to the formation of (100) and (111) internal surfaces above 400 °C via agglomeration of the initial defect structure. In addition, molecular hydrogen is evolved between 200 and 400 °C and subsequently traps in the microvoids bounded by the internal surfaces, resulting in the build-up of internal pressure. This, in turn, leads to the observed "blistering" of unconstrained silicon samples, or complete layer transfer for silicon wafers joined to a supporting (handle) wafer which acts as a mechanical "stiffener." © 1997 American Vacuum Society. [S0734-211X(97)08904-X]

I. INTRODUCTION

The physics and chemistry of hydrogen in silicon has been the subject of considerable scientific and technological interest for over three decades. This interest has been driven, in part, by the ubiquitous occurrence of hydrogen in semiconductor processing which invariably leads to incorporation into the substrate either intentionally or unintentionally. Importantly, this hydrogen may profoundly alter the electrical characteristics of the resultant device by diffusion into the active region and passivation of the dopant. The myriad studies of hydrogen that have resulted are summarized in a number of review texts, to which the reader is referred for a comprehensive treatise.¹⁻³

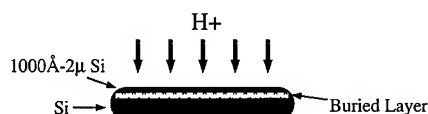
Recently, there has been renewed focus on the internal transformations that occur upon implantation of hydrogen into semiconductors due to the demonstration of the hydrogen-induced shearing of macroscopic layers of silicon, with complete transfer to a supporting wafer to which the implanted material had previously been joined.⁴ In this way, it was found that thin (1000 Å – 2 μm), crystalline films of Si could be formed on top of an SiO₂ layer which was itself bonded to another Si wafer (the support or "handle"). That is to say that extremely high-quality silicon-on-insulator (SOI) structures could be formed directly during annealing (as schematically illustrated in Fig. 1), with vastly improved thickness uniformity [total thickness variation (TTV) = ±50 Å over a 100 mm wafer] relative to that obtained using conventional wafer bonding technology, for which extensive grinding and etching is required to form the final, thin silicon layer.⁵ This remarkable process has also subsequently been applied to the transfer of thin layers of exotic materials (e.g., SiC, InP) onto supporting wafers and therefore shows real promise as a generally applicable material synthesis protocol.

Despite the potential importance of this process, the fundamental phenomena that drive the layer shearing and dictate the quality of the transferred layer, remain poorly understood. For example, it is not known whether the principal role of hydrogen is "chemical," i.e., it passivates the internal Si dangling bonds preventing reformation of Si-Si linkages or facilitates bond scission by concomitant insertion into the Si-Si bonds. In this case, the relative kinetics/thermodynamics of the Si-H species formed will strongly influence the internal structure that develops upon annealing. Alternatively, the dominant process may be "physical/mechanical" in nature, in that the creation of internal pressure by H₂ formation, combined with the damage caused by the implanted hydrogen may effectively cause fracture of the Si substrate. If so, the recombination kinetics of dihydrogen formation (which may, in turn, depend on the transport or diffusion of H from point defects), or the development of appropriate regions of free space (i.e., voids) in which to evolve the H₂ gas, may dictate the rate of exfoliation. Clearly, much remains to be resolved at the microscopic and macroscopic level.

Therefore, we have undertaken a wide-ranging series of studies of the thermal evolution of hydrogen implanted into silicon, as a function of implantation dose, depth, and annealing temperature, the results of which are summarized in this article. Using a combination of infrared (IR) spectroscopy, forward recoil scattering (FRS), transmission electron microscopy (TEM), and atomic force microscopy (AFM), we have been able to elucidate the key transformations that occur in both nonbonded and bonded silicon substrates. Specifically, we find that the mechanism consists of the following essential components (i) Initial formation of hydrogenated point defects complexes in the regions of lower hydrogen concentration (i.e., at the periphery of the implanted region), as well as extensive disruption of the silicon lattice (due to multivacancy formation) and 50–100 Å platelet formation near the peak of the implantation profile;

^{a)}Solid State Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831.

1) Ion implantation



2) Invert, Bond and Anneal

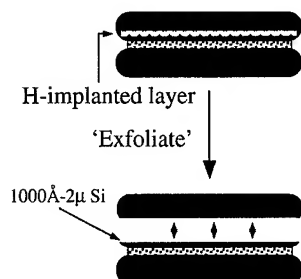


FIG. 1. Schematic illustration of the commercial SmartCut™ process, in which an oxidized Si wafer is preimplanted with H^+ , then subsequently joined to a second (handle) wafer and annealed, resulting in the direct formation of a high quality SOI substrate.

(ii) Collapse of this defective, multivacancy structure resulting in formation of H_2 and agglomeration of the remaining bound hydrogen into vacancy defect complexes such as $VH_{3,4}$ upon annealing to 400 °C; (iii) Rearrangement of the defect structure above 400 °C into extended internal hydrogen-terminated (100) and (111) surfaces; (iv) Trapping of H_2 in the microvoids bounded by the internal surfaces, resulting in the build-up of internal pressure; (v) Development of these microvoids into macroscopic (pressurized) cracks that either lead to blistering of unconstrained silicon samples, or complete layer transfer for silicon wafers joined to a supporting (handle) wafer. We show that the role of the handle wafer is to provide the necessary restoring force to drive crack propagation parallel to the surface, effectively suppressing the vertical lift off that leads to blistering.

II. EXPERIMENT

All studies were performed on Si(100) wafers [100 mm FZ, boron doped (*p* type), resistivity 5–10 Ω cm] implanted with H^+ (through a 3000 Å layer of thermally grown oxide) at doses ranging from 2×10^{16} to 1.8×10^{17} H/cm² and energies ranging from 75 keV to 1 MeV, resulting in implantation depths of 3300 Å to 16 μ m. The substrate was maintained at close to ambient temperature throughout the implantation. The samples were given a modified RCA clean prior to analysis to remove surface contaminants introduced by the implantation process. Finally, some of the implanted samples were joined to an unimplanted handle wafer that was terminated by the thin chemical oxide layer formed by the RCA clean.^{6,7} All wafers were cut into smaller samples for analysis. Annealing was performed sequentially for 30 min at each temperature in a tube furnace either under vacuum ($P \sim 2 \times 10^{-8}$ Torr) or a positive pressure of N_2 . No

dependence on the anneal ambient (vacuum or N_2) was found, but the shearing/layer transfer process was found to be a sensitive function of the annealing protocol and thermal history of the sample.

The infrared data were obtained using a Nicolet 60 SX FTIR spectrometer equipped with a broadband MCT detector and operating at 2 cm⁻¹ resolution in a multiple internal reflection (MIR) geometry. All spectra were collected under ambient conditions and referenced either to an identical unimplanted sample or an implanted sample that had previously been annealed to 1100 °C, at which temperature no hydrogen remains. Additional baseline subtraction was usually required to remove oscillations that result from interference between beams reflected from the implanted region (after partial separation/crack development) and the sample surface. In each case, this subtraction was performed using the same high order polynomial function that was fit to the entire broadband spectrum (1500–8000 cm⁻¹) to minimize the preferential selection of spectral features.

Samples for cross-sectional imaging by transmission electron microscopy (TEM) were prepared in the conventional way by mechanical thinning and Ar^+ ion beam milling. The H implantation-induced damage and defects were imaged in bright field and dark field conditions using the weakly excited 400 reflection, and in axial [110] high resolution.

Forward recoil scattering studies were performed using 2.6 MeV He^+ ions incident at 15° to the surface plane and recoiled H was detected at 30° to the incident beam direction. Concentration calibration was performed by comparison to a standard sample implanted with a known hydrogen concentration.

Atomic force microscopy (AFM) was performed using a Nanoscope III (Digital Instruments, Santa Barbara, CA) in either contact or tapping mode. Large-scale images ($\geq 25 \mu$ m \times 25 μ m) were obtained in contact mode with silicon nitride tips. Roughness values of the exfoliated and substrate surfaces were determined using a tapping mode with etched silicon tips. The reported roughness value is the average of six measurements made at different scan sizes ranging from 0.1 to 5 μ m.

III. RESULTS AND INTERPRETATION

A. Infrared spectroscopy

Representative spectra for both a low dose (2×10^{16} H/cm²) and an intermediate dose (6×10^{16} H/cm²) of hydrogen implanted into Si(100) are shown in Fig. 2, as a function of annealing temperature. For the as-implanted low dose sample, the 1800–2300 cm⁻¹ region consists of at least 11 discrete Si–H stretching modes, superimposed on a broadband centered at ~ 2000 cm⁻¹. Similar features are observed for the intermediate dose; however, the intensity of broadband absorption is increased markedly, so that it dominates the discrete mode structure. The origin of these two types of spectral features (that are common to all H^+ doses that we have studied) can be assigned in a qualitative sense by reference to the extensive literature. Although the precise

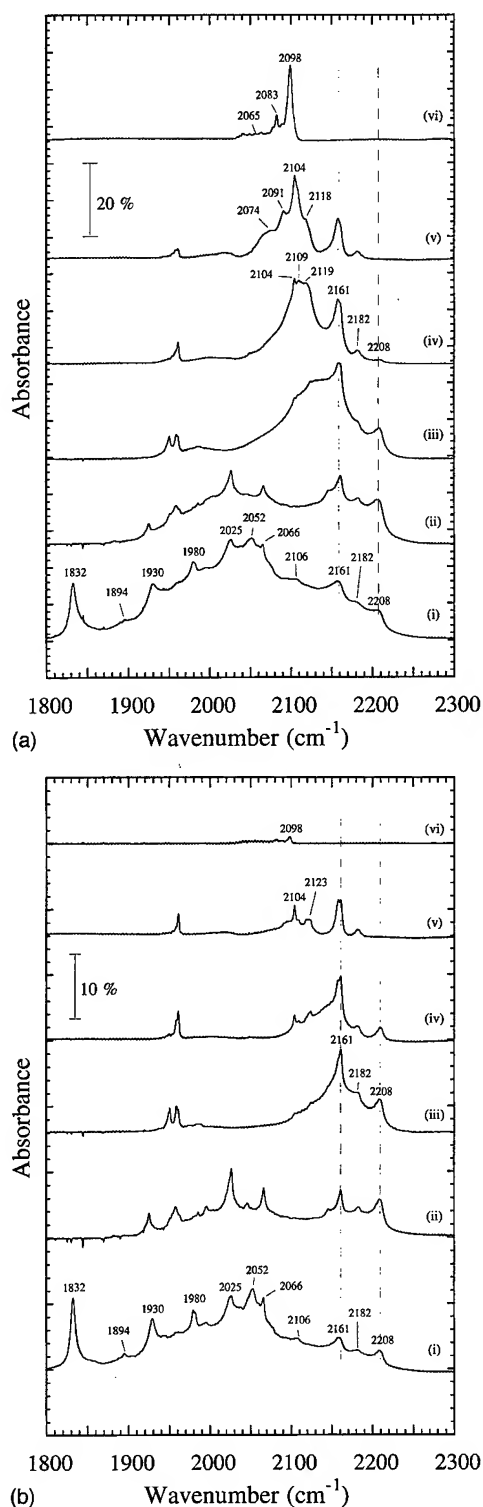


FIG. 2. Infrared spectra of the Si-H stretching modes of hydrogen implanted 3300 Å into Si(100) at doses of (a) 6×10^{16} H/cm² (intermediate dose) and (b) 2×10^{16} H/cm² (low dose) at (i) room temperature and after annealing to; (ii) 300 °C; (iii) 425 °C; (iv) 500 °C; (v) 550 °C; and (vi) 650 °C.

assignment of each specific feature is still the subject of considerable controversy,¹ there is general consensus that the discrete modes are associated with isolated point defect type complexes of the form V_xH_y or Si_lH_y (where "V" is used to

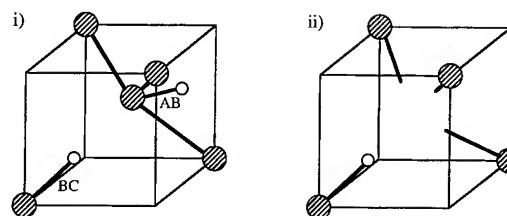


FIG. 3. Illustration of (local) Si-H defects observed in this work; (i) H_2^* defect and; (ii) Monovacancy monohydride VH ($VH_{3,4}$ have 3 and 4 of the dangling bonds of the monovacancy terminated with H, respectively). Larger shaded circles=Si; smaller open circles=H.

denote a Si vacancy, "I" denotes a Si interstitial, and the subscripts may have the values $x=1$ and 2 and $y=1-4$), imbedded in crystalline silicon, whereas the broad absorption is due to H trapped in a highly defective Si region. In fact, this broadband absorption is strongly reminiscent of that observed for hydrogenated amorphous silicon, $a\text{-Si-H}$, which has been attributed to a inhomogeneous distribution of monohydride-terminated, multivacancy defects also of the form V_xH_y , but with $x=3$ or 4 and $y=1$.¹ In accordance with this assignment, we ascribe the band centered at 2000 cm^{-1} to such a "multivacancy" defect region.

From our studies of a wide variety of different H^+ doses and depth distributions, we find that the point defect mode intensity is largely invariant with dose, whereas the multivacancy signature scales with the hydrogen concentration, so that it is probable that the point defects occur at the periphery of the multivacancy region which, in turn, occurs around the peak of the hydrogen concentration profile. The specific assignments of some of these point defect modes are still somewhat controversial; however, it is instructive to consider the nature of the modes for which more definitive assignments can be made, before continuing to discuss the annealing data.

The pair of modes observed at 1832 and 2052 cm^{-1} in the as-implanted samples are assigned to the H_2^* defect [Fig. 3(a)], by reference to the detailed studies of this species by Bech Nielsen *et al.*³ Specifically, the two Si-H modes are due the (weakly) coupled motions of a hydrogen atom bound close to the bond center (BC) site in the Si lattice (2052 cm^{-1}) and the other at the antibonding (AB) site (1832 cm^{-1}). Importantly, the close agreement between the frequencies and annealing behavior ($T < 250$ °C) observed in this work, with those previously reported in the literature (1838/2062 cm^{-1} and $T < 250$ °C), strongly support this assignment.

In addition, the modes observed above 2050 cm^{-1} are primarily assigned to hydrogen bound at monovacancy defects of the form VH_y where $y=1-4$ [Fig. 3(b)], by reference to the literature. Recent isotopic labeling studies, in combination with uniaxial stress measurements of proton-implanted Si have afforded the following assignments:⁶ The feature at 2066 cm^{-1} is assigned to a single hydrogen bound to the monovacancy defect (VH); the features at 2161 and

2182 cm^{-1} are due to three hydrogen atoms bound to the monovacancy (VH_3), whereas the 2208 cm^{-1} band is due to VH_4 .⁷ The feature at 2106 cm^{-1} is intriguing as it occurs in the frequency range usually ascribed to hydrogen bound to extended Si surfaces (see the following). Therefore, given that TEM images of the as-implanted sample show (111) and (100) "seed" platelets (Sec. II), we tentatively assign this feature to hydrogen terminating these 50–100 Å platelets. Finally, the bands at 1930 and 2025 cm^{-1} are assigned to small divacancy clusters terminated by either one or two hydrogens, i.e., V_2H or V_2H_2 .¹ Now, considerable progress can be made in understanding the broad transformations that occur upon annealing, based on these assignments and our recent definitive proof of the origin of modes that appear at $\sim 2100 \text{ cm}^{-1}$ at higher temperatures,⁸ as summarized below.

There is a general attenuation of the broadband, multivacancy structure after annealing to 300 °C, as well as the complete loss of the H_2^* modes at 1832 and 2052 cm^{-1} . In contrast, the VH_3 and VH_4 multiply-hydrogenated monovacancy modes at 2161/2182 and 2208 cm^{-1} , respectively, are all enhanced relative to the as-implanted spectrum. Based on these observations, we can immediately conclude that there is both a net loss of *bound* hydrogen and agglomeration of the remaining trapped hydrogen at existing vacancies. These initial trends are exacerbated after annealing to 425 °C for both samples [Figs. 2(a)(iii) and 2(b)(iii)]. Essentially complete attenuation of the multivacancy signature occurs, together with a pronounced growth of the agglomerated hydrogen monovacancy modes discussed above, so that the latter modes dominate the observed spectrum. After annealing to higher temperatures, the similarity in the evolution of the features for the low and intermediate H^+ doses disappears. Most notably, for the $6 \times 10^{16} \text{ H/cm}^2$ dose the monotonic loss of the VH_3 and VH_4 modes is accompanied by the growth of discrete spectral features at 2100–2120 cm^{-1} , which we will show have profound mechanistic implications. In contrast, the spectra of the $2 \times 10^{16} \text{ H/cm}^2$ dose simply show attenuation of the agglomerated monovacancy modes, with almost no attendant growth of new spectral features.

Specifically, new modes appear at 2104, 2109, and 2119 cm^{-1} in the spectrum of an intermediate hydrogen dose annealed to 500 °C [Fig. 2(a)(iv)]. Upon further annealing to 550 °C, the two higher frequency peaks collapse, leaving only the sharp 2104 cm^{-1} feature [Fig. 2(a)(v)] with a distinct shoulder at 2091 cm^{-1} and a broader band centered at 2074 cm^{-1} . Finally, at 650 °C, the 2104 cm^{-1} mode is replaced by an even sharper peak at 2098 cm^{-1} and a new discrete feature is apparent at 2083 cm^{-1} [Fig. 2(a)(vi)]. The assignment of these features is central to the understanding of the exfoliation process, since they develop coincident with the appearance of surface blistering in these (nonbonded) samples. More importantly, this blistering is not observed for the low dose, $2 \times 10^{16} \text{ H/cm}^2$ sample for which these vibrational modes are demonstrably absent. Conversely, the appearance of modes at $\sim 2100 \text{ cm}^{-1}$ occurs at lower temperatures for doses of 1.0×10^{17} and 1.8×10^{17} , for which the

onset of blistering occurs at 450 and 400 °C, respectively (data not shown).

We propose that the modes at 2060–2120 cm^{-1} are due to hydrogen trapped at extended internal surfaces, based on the remarkable agreement with the known surface literature and extensive additional studies of internal hydrogen-terminated surfaces formed by wafer bonding.^{9,10} A rigorous discussion of the assignment of these modes is given elsewhere,⁸ the essence of which is summarized as follows: The bands at 2109 and 2119 cm^{-1} are due to the symmetric and asymmetric Si– H_2 stretching motions of atomically rough (100) surfaces, based on the excellent agreement with the frequencies reported for dilute HF etched Si(100), for which $\nu_{\text{ss}}(\text{SiH}_2)$ and $\nu_{\text{as}}(\text{SiH}_2)$ are observed at 2109 and 2120 cm^{-1} , respectively.^{11,12} The mode at 2104 cm^{-1} present at 500–550 °C is paired with the mode at 2091 cm^{-1} that is clearly resolved in the 550 °C spectrum and attributed to $\nu_{\text{as}}(\text{SiH}_2)$ and $\nu_{\text{ss}}(\text{SiH}_2)$ of atomically smoother dihydride terminated Si(100) surfaces.¹³ Once again, there is precise agreement with the literature values of 2103 and 2091 cm^{-1} for dihydride prepared by UHV exposure of Si(100) to atomic hydrogen. These assignments essentially dictate that, upon further annealing, these internal surfaces should undergo the (2×1) reconstruction to produce the coupled monohydride species, H–Si–Si–H, just as is observed for hydrogen-terminated Si(100) surfaces in UHV. Indeed, this is precisely what is observed; the modes at 2098 and 2074 cm^{-1} , observed above 550 °C, are characteristic¹⁴ of coupled monohydride species on the atomically smooth terraces and along the steps of a Si(100) surface, respectively!

Finally, the sharp mode at 2083 cm^{-1} in the spectrum of $6 \times 10^{16} \text{ H/cm}^2$ annealed to 650 °C, can be definitively assigned to an atomically flat Si(111) surface.¹⁵ In fact, the weak broadband structure observed at lower frequency ($\sim 2065 \text{ cm}^{-1}$) has recently been assigned to the coupled mode formed by two opposing, *closely interacting* H-(1 \times 1)Si(111) surfaces.¹⁰ Both theoretical and experimental studies of the bonding of two such surfaces clearly demonstrate that when the intersurface separation approaches $\sim 1\text{--}2 \text{ \AA}$, the van der Waals interaction gives rise to a perturbed mode at 2065 cm^{-1} , in addition to the unperturbed mode at 2083 cm^{-1} . Finally, we have recently obtained further confirmation of the internal surface character of these modes by studying the low frequency, bending modes ($\sim 615\text{--}650 \text{ cm}^{-1}$) associated with these Si–H stretching frequencies⁸ and we find that, once again, there is excellent agreement with the data obtained for hydrogen-terminated Si(111) and (100) surfaces.

B. Transmission electron microscopy

We have also studied the thermal evolution of the implanted region using cross-sectional TEM, in order to establish the existence of internal (100) and (111) extended defects. In Fig. 4, a typical electron micrograph of the as-implanted region is shown, together with an image obtained after annealing a sample implanted with $1 \times 10^{17} \text{ H/cm}^2$ to

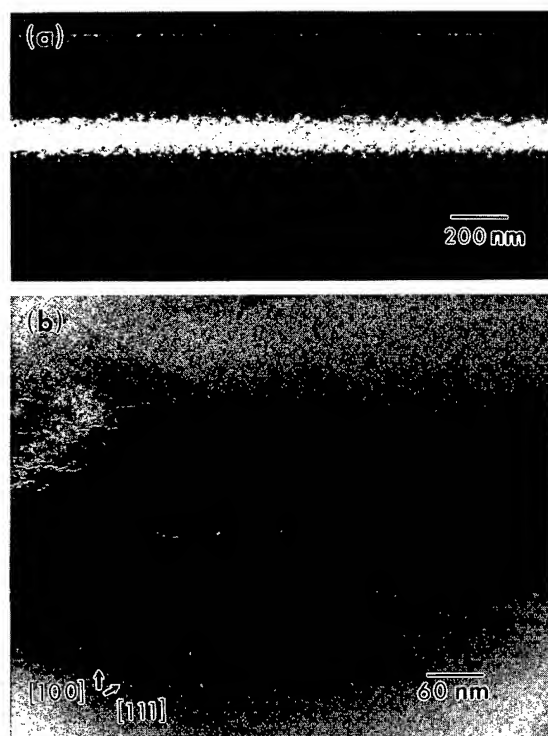


FIG. 4. Cross-sectional TEM images of Si(100) implanted with 1×10^{17} H/cm² after; (a) room temperature implantation and; (b) annealing to 475 °C [image obtained with higher magnification than (a)].

475 °C. The as-implanted image shows a highly defective (but not truly “amorphous”) layer, located approximately 2800–4300 Å below the surface. Both bright field and high resolution microscopy of a thinner section of the TEM foil (image not shown) reveal that the defective layer contains some plateletlike defects (50–100 Å) oriented along the (100) and (111) planes, as has been previously observed for silicon exposed to hydrogen-containing plasmas.¹⁶ Interestingly, after annealing to 475 °C, a connected network of microcracks appears [Fig. 4(b)] forming a corrugated “macro” crack (2–3 μm in length) parallel to the (100) surface. These cracks are seen to consist predominantly of (100) planes that are connected by a minority of (111) planes, in excellent agreement with our conclusions based on infrared data.

C. Optical and atomic force microscopy

The effect of the internal crack formation on the overlying silicon has been investigated using both optical microscopy and AFM (Fig. 5). The initially pristine surface is visibly disrupted on annealing a sample implanted with 1×10^{17} H/cm² to 475 °C. The optical micrograph clearly shows the surface consists of both “popped” blisters with sizes ranging from 10 to 250 μm, and “capped” bubbles of dimension 10–70 μm. Additional information is available in the atomic force microscope image of the same sample [Fig. 5(b)]. The depth of the crater formed by this blistering can be measured and is found to be equal to the implantation depth, i.e., the peak of the implanted H concentration profile, where the crack formation is most pronounced. This observation

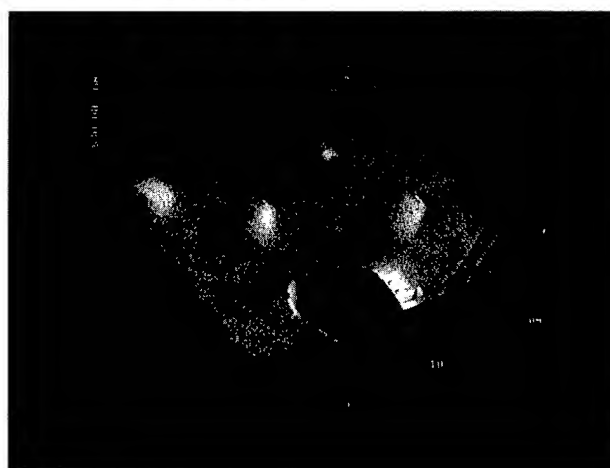
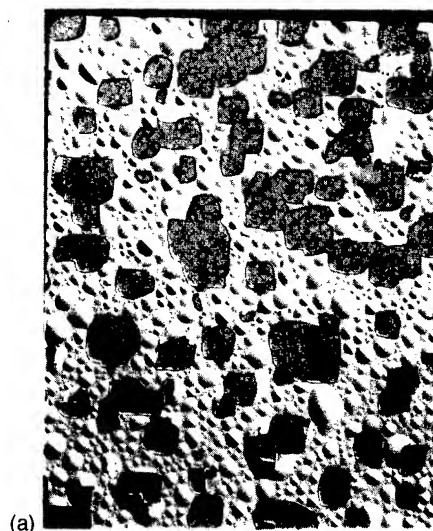


FIG. 5. (a) Optical micrograph of 1×10^{17} H/cm² after annealing to 475 °C (Mag. $\times 90$). (b) AFM image of a 15 μm \times 15 μm region of the same sample as in (a).

confirms the intuitive notion that the underlying cracks lead to the lift off of the overlying Si. More importantly, the presence of the capped bubbles on the Si surface strongly suggests that there is an internal pressure that causes elastic deformation of the overlying Si. In fact, using the elastic modulus of Si and assuming a bubble of dimensions 100 Å (height) \times 2 μm (diameter) as observed by AFM, we estimate that the internal pressure must approach ~ 10 kbar, *even for capped bubbles* which, presumably, are not sufficiently pressurized to give rise to blistering. Finally, the rms roughness measured at the bottom of the crater is found to be only 80 Å, which underlines the power of this approach for high quality layer transfer.

D. Forward recoil scattering

We have employed FRS to complete the arsenal of probes used to study these samples. Importantly, it is a probe of the total hydrogen content, in contrast to IR spectroscopy which is not sensitive to the presence of molecular H₂ under normal

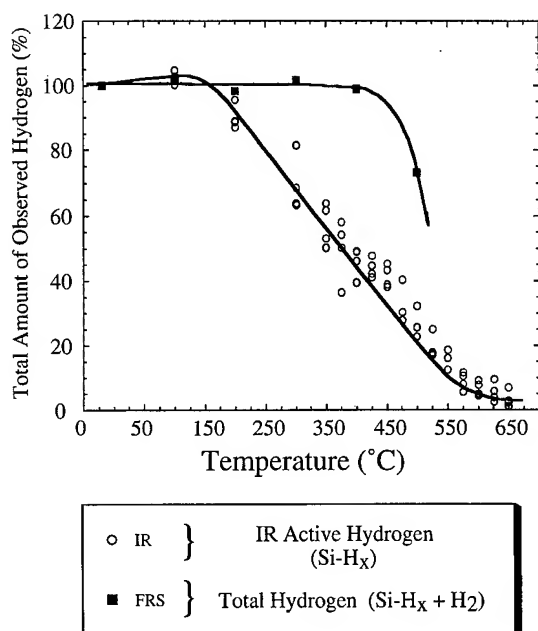


FIG. 6. Plot of the integrated hydrogen signal obtained using forward recoil scattering and IR spectroscopy.

conditions. The total amount of hydrogen, determined by FRS, is plotted as a function of annealing temperature in Fig. 6. In addition, the normalized integrated IR absorbance for each hydrogen dose is also plotted on the same axes. The disparity in the temperature dependence is immediately apparent on comparison of the data obtained using the two techniques; the amount of bound hydrogen monotonically decreases above 200 °C, while the total hydrogen content remains constant up to 400 °C. Clearly, there has been a conversion of trapped hydrogen to an unbound form, i.e., either atomic or molecular hydrogen. The existence of isolated H atoms can be effectively eliminated since it is highly improbable that such a state would be stable with respect to trapping at a vacancy or interstitial defect upon cooling to room temperature (where the data were acquired). Therefore, the only reasonable interpretation is that there is net conversion of Si-H into H₂ at temperatures between 200 and 500 °C, and that the molecular hydrogen remains trapped in the Si crystal. Interestingly, previous theoretical investigations of hydrogen in intrinsic Si found that H₂ is indeed stable at the tetrahedral interstitial sites.¹⁷ However, the solubility of H₂ in room temperature silicon is vanishingly small.³ In addition, the diffusion length of hydrogen in Si is greater than the implantation depth (3300 Å) over this temperature range,¹ so that the molecular hydrogen cannot be trapped in the Si lattice *per se*. We propose that the H₂ must be bound at internal voids or cavities, as has been observed for hydrogenated amorphous silicon (*a*-Si:H). Notably, molecular hydrogen was found to accumulate in the 10–20 Å pores in *a*-Si:H, for temperatures up to 400 °C, resulting in internal pressures as high as 100 kBar.¹⁸ These voids are distinct from the hydrogenated VH_{3,4} defects discussed above, since these monovacancy defects cannot accommo-

date a H₂ molecule. Therefore, it is probable that hydrogen traps voids of the type V_x (where $x > 1$) that are produced from the initial multivacancy structure. Indeed, since these voids are not hydrogen terminated (based on the absence of any related Si-H modes observed between 200 and 350 °C), it is plausible that the H₂ is produced by recombination of the Si-H monohydride species that initially decorate the multivacancy structures, and traps locally in the hydrogen-free multivacancies so formed. In this picture, the coalescence of the initial multivacancy structure to form the agglomerated monovacancy VH_{3,4} defects is a competing reaction pathway.

The last point of note is that the rate of conversion of Si-H to H₂ is equivalent for both low (2×10^{16} H/cm²) and high hydrogen doses (1.0×10^{17} H/cm²), i.e., the FRS and IR data show identical functional dependencies independent of dose. Yet, the samples implanted with 2×10^{16} H/cm² outwardly show a clear difference in their annealing behavior in that they do not give rise to exfoliation. This observation has important mechanistic implications that are discussed below.

IV. DISCUSSION

In the preceding section, the essential elements of the exfoliation process were highlighted and discussed separately. In the remainder of this article, we will combine these elements to propose a coherent mechanism that has phenomenological utility.

The first important observation is that it is the “multivacancy” defect structure that produces molecular hydrogen on annealing up to 350 °C, by comparison of the integrated FRS and IR data. Specifically, the total (crystalline) point defect absorption is constant up to 350 °C, in contrast to the ~40% attenuation of the broadband component of the IR absorption over this temperature range. Subsequently, on annealing to ~400 °C, attenuation of all the vibrational features below 2100 cm⁻¹ occurs, with the continued production of H₂ as well as the formation of the agglomerated VH_{3,4} type defects. Both the H₂ production and VH_{3,4} defect formation scale with the hydrogen dose, which necessitates that they are predominantly derived from the multivacancy region (see earlier). Furthermore, previous studies of lower hydrogen doses (1×10^{16} H/cm²), which do not exhibit any appreciable broadband absorption at 2000 cm⁻¹, found that no IR inactive species were formed at any temperature, but reported pronounced VH_{3,4} features at 2160 and 2210 cm⁻¹.¹⁹ Therefore we conclude that only the multivacancy region yields H₂ in our studies, but that both defect regions support the growth of the agglomerated defects.

Secondly, we propose that the agglomerated VH_{3,4} defects are precursors to the formation of extended internal surfaces, based on the attenuation of the former *concomitant* with the appearance of the latter. Moreover, it is reasonable to postulate that the initial 50–100 Å (100) and (111) platelets observed following implantation may serve to seed or nucleate the formation of the extended surfaces, by attachment of these agglomerated defect structures. In either case, the proposed role of VH_{3,4} defects implies that the kinetics of inter-

nal surface formation should be a function of the density of such defects. Indeed, this is exactly what is observed experimentally; the rate of crack propagation increases with the integrated absorbance of the 2161, 2182, and 2208 cm^{-1} bands. Furthermore, there is a minimum density of $\text{VH}_{3,4}$ defects, below which efficient extended surface formation does not occur, as is manifestly the case for the low ($2 \times 10^{16} \text{ H}^+/\text{cm}^2$) dose.

Now, the question arises as to the exact role that the H_2 formation plays in the exfoliation process. We have already intimated that an internal pressure must develop in the implanted region, by observation of pressurized capped bubbles for intermediate to high hydrogen doses (Fig. 5), and that internal cracks are somehow involved, but now we will further substantiate the causal link between these observations.

The IR and TEM data clearly demonstrate the presence of the internal (100) and (111) surfaces after annealing samples implanted with 6×10^{16} – $1.8 \times 10^{17} \text{ H}^+/\text{cm}^2$ to between 350 and 450 °C. Furthermore, we find that exfoliation only occurs after the appearance of these internal surfaces/cracks and *does not occur* when the crack density is too low, i.e., for H doses of $2 \times 10^{16} \text{ H}^+/\text{cm}^2$. In addition, the blistering is found to originate in the regions of the highest crack density, which strongly suggests that the pressure is developing in the cracks. All that remains is to prove that H_2 is indeed present and is the source of the internal pressure. This we have verified using *in situ* mass spectrometry during the annealing process. Specifically, we observed a massive burst of molecular hydrogen coincident with the occurrence of exfoliation, with no other gaseous products detected at any temperature. Interestingly, by careful calibration of the mass spectrometer, we were also able to quantify the total amount of H_2 evolved and found that only 30% of the initial H implant is liberated in this step; IR spectra of the two sheared pieces clearly demonstrate that a substantial amount of hydrogen remains trapped at internal cracks that persist beneath the newly formed surfaces (data not shown).

So, a coherent mechanism is beginning to emerge: At sufficiently high doses, implantation of hydrogen into silicon predominantly causes the formation of multivacancy structure. Upon annealing to temperatures between 200 and 350 °C, this multivacancy defect region simultaneously liberates H_2 and partially evolves into a small-cavity structure, which traps the H_2 so formed. In addition, agglomerated monovacancy defects are also produced from the initial multivacancy structure which, in turn, coalesce (or nucleate at the 50–100 Å platelets) to form internal (100) and (111) surfaces. The molecular hydrogen subsequently accumulates in the microscopic cracks that are defined by these extended surfaces, resulting in the development of sufficiently high internal pressures to cause the exfoliation of the overlying silicon. So, the role of the implanted H is twofold; (i) It acts “chemically” in that it drives the formation of microscopically flat internal surfaces and acts as a source of gaseous H_2 which traps in the internal cavities where; (ii) it acts “physically” as an internal pressure source.

The preceding mechanism suggests that a critical kinetic

regime exists, wherein the molecular H_2 is liberated from the (angstrom scale) voids and must retrap in the incipient microcracks, in order to prevent diffusion away from the implanted region and allow the creation of the requisite internal pressure. Clear evidence for this is provided by reference by the data obtained for low H doses ($2 \times 10^{16} \text{ H}^+/\text{cm}^2$), for which the crack density is found to be very low, so that the molecular hydrogen can, by the preceding analysis, escape from the implanted region and diffuse into the bulk crystalline silicon. This is indeed what is observed for such a low H dose sample when it is prejoined to a handle wafer. Specifically, the out-diffusing H_2 is found to accumulate in the microvoids present at the *joined* interface, upon annealing to 550 °C, resulting in the build-up of pressure and eventual separation at this interface.

These mechanistic deductions also naturally explain the slightly surprising observation that the thermal evolution of the integrated FRS/IR data are similar for both low and high H^+ doses, despite the vastly different crack densities observed in the two cases. This *apparent* similarity in the evolution of the implanted hydrogen, results from the similar kinetics for hydrogen out diffusion (which occurs for the $2 \times 10^{16} \text{ H}^+/\text{cm}^2$ dose) and internal surface formation/blistering (that is dominant for the $1 \times 10^{17} \text{ H}^+/\text{cm}^2$ dose); both of which directly result in a net loss of hydrogen on annealing to 450 °C.

Clearly, a number of key processes, i.e., crack formation, hydrogen diffusion, and H_2 trapping in voids, are all operative in the same temperature regime, but the question remains as to which is the *rate-limiting* step. In order to address this, we have performed a series of annealing experiments in which the time to induce exfoliation was measured as a function of different annealing temperatures for a dose of $6 \times 10^{16} \text{ H}^+/\text{cm}^2$. Then a simple Arrhenius analysis was performed to yield an “effective” activation energy of 1.8 eV. Interestingly, it has recently been shown that the values obtained by such an analysis are in reasonably good agreement with the bond energy of the implanted material for a variety of different substrates such as Si, SiC, and diamond,²⁰ suggesting that the *rate-limiting step is the rupture of the remaining Si–Si bonds* in the implanted region. Indeed, the value that we obtain is close to the value of 2.3 eV for the bulk Si–Si bond energy. The fact that the value is actually ~20% lower than the nominal Si–Si bond energy is not surprising since the thermal heating serves not only to activate Si–Si bonds (as assumed in the Arrhenius analysis) but also to increase the pressure of the gas in the cracks which is not explicitly accounted for. Therefore, the value so derived should be lower than the material bond energy because the additional pressure-induced activation will reduce the value of T required to drive the exfoliation. However, this does not preclude the contribution of other physical and chemical processes; e.g., the simultaneous insertion of H into the highly stressed Si–Si bonds at the boundaries of the microcracks would also facilitate macroscopic crack propagation and lead to a consequent reduction in the measured activation energy. Clearly, a more complete treatment is

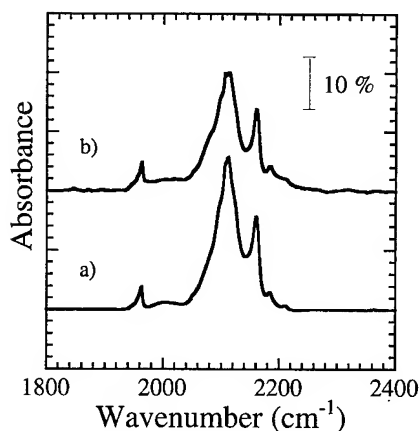


FIG. 7. Comparison of IR spectra of Si(100) implanted with 6×10^{16} H/cm² to a depth of 3300 Å after annealing to 475 °C for 20 min; (a) nonbonded and (b) bonded sample.

required before a quantitative understanding can be claimed.

In the final section of this article, we will discuss the transition from blistering to complete layer transfer that occurs upon joining an implanted sample to another, handle wafer. Given the drastic nature of the change that occurs, i.e., shearing of complete 100–200 μm layers instead of formation of ~ 2 –250 μm blisters, one might expect a significant enhancement in the development of cracks in the bonded wafer pair. In order to investigate this aspect of the process, we have annealed both bonded and nonbonded samples under identical conditions and studied the internal surface formation using both IR and TEM. Surprisingly, we find that there is essentially no difference observed between the two samples using either technique, even immediately prior to the onset of exfoliation (Figs. 7 and 8). This suggests that the effect of the bonded wafer is dynamic, in that it is only observed as the blistering/layer transfer process is occurring. Given this, we conclude that the role of the handle wafer is to provide a mechanical restoring force that opposes the vertical lift off that leads to blistering and drives lateral crack propagation instead, as schematically illustrated in Fig. 9.

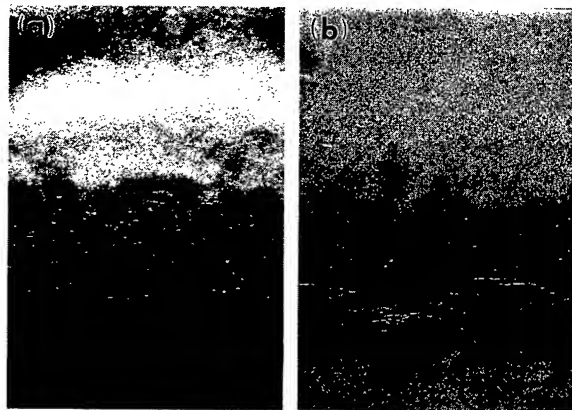


FIG. 8. Comparison of TEM images of Si(100) implanted with 6×10^{16} H/cm² to a depth of 3300 Å after annealing to 475 °C for 20 min; (a) nonbonded and (b) bonded sample.

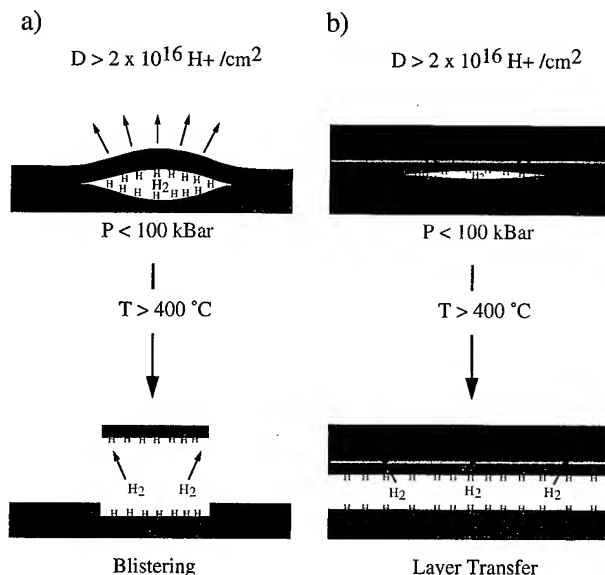


FIG. 9. Schematic illustration of (a) the blistering step and (b) the mechanical action of the handle wafer that leads to the lateral propagation of cracks.

Verification of the mechanical role of the support is provided by the observation that samples implanted with hydrogen at a depth of 16 μm, show similar large-scale layer exfoliation even in the absence of the bonded “stiffener.”

In summary, we have presented an overview of the extensive studies we have undertaken of the physical and chemical processes intrinsic to the hydrogen-induced exfoliation of silicon. By using a combination of techniques and a wide range of different experimental conditions, a detailed (qualitative) mechanism has begun to emerge, although much remains to be done to definitively characterize all the essential elements of this highly complex process. In particular, the predominance of the (100) internal surfaces relative to (111) and the mechanism and kinetics of the hydrogen diffusion through the implanted region need to be substantiated in order to understand the extended crack propagation and the observation that only 30% of the hydrogen contributes to the internal pressure. To this end, we are currently investigating the effect of implantation depth and annealing protocol, as well as the coimplantation of hydrogen and helium, on the evolution of the implanted region. However, many additional experimental and theoretical studies of the internal chemistry, physics, and mechanics of this process are required before a complete quantitative mechanism emerges.

ACKNOWLEDGMENTS

The authors would like to thank Krishnan Raghavachari and Boris Stefanov for many stimulating discussions and C. A. Goodwin and C.-M. Hsieh for financial and technical support. Y. C. acknowledges a grant from the Belgian Fund for Industrial and Agricultural Research (F.R.I.A) and a travel grant from the Belgian French speaking community.

¹S. J. Pearton, J. W. Corbett, and M. Stavola, *Hydrogen in Crystalline Semiconductors* (Springer, Berlin, 1992), and references therein.

- ²*Hydrogen in Semiconductors*, edited by J. I. Pankove and N. M. Johnson (Academic, New York, 1990).
- ³*Defects in Semiconductors 17, Materials Science Forum*, edited by H. Heinrich and W. Jantsch (Trans Tech, Switzerland, 1993), pp. 143–147.
- ⁴A. J. Auberton-Hervé, J. M. Lamure, T. Barge, M. Bruel, B. Aspar, and J. L. Pelloie, *Silicon-On-Insulator Technology*, Semi-Con West, 1995.
- ⁵*Proceedings of the 1st, 2nd, and 3rd International Symposia on Semiconductor Wafer Bonding: Science, Technology and Applications* (The Electrochemical Society, Pennington, NJ, 1992–1995).
- ⁶B. Bech Nielsen, L. Hoffmann, and M. Budde, *Mater. Sci. Eng. B* **36**, 259 (1996).
- ⁷The modes above 2050 cm^{-1} had been previously assigned to either VH_2 or Si_2H_2 species (Ref. 1). However, following the recent definitive work of Bech Nielsen *et al.* (Ref. 8), all of these features can be attributed to VH_2 species, exclusively. Therefore, the fate of the interstitials that are necessarily produced by vacancy creation is, as yet, unresolved.
- ⁸V. E. Marsico, M. K. Weldon, Y. J. Chabal, Y. Caudano, A. Agarwal, and D. J. Eaglesham (unpublished).
- ⁹Y. J. Chabal, M. A. Hines, and D. Feijoo, *J. Vac. Sci. Technol. A* **13**, 1719 (1995); D. Feijoo, Y. J. Chabal, and S. B. Christman, *Appl. Phys. Lett.* **65**, 2548 (1994).
- ¹⁰M. K. Weldon, Y. J. Chabal, D. R. Hamann, S. B. Christman, E. E. Chaban, and L. C. Feldman, *J. Vac. Sci. Technol. B* **14**, 3095 (1996); M. K. Weldon, V. E. Marsico, Y. J. Chabal, D. R. Hamann, S. B. Christman, and E. E. Chaban, *Surf. Sci.* **368**, 163 (1996).
- ¹¹P. Dumas, Y. J. Chabal, and P. Jacob, *Surf. Sci.* **269/270**, 867 (1992).
- ¹²Y. J. Chabal, G. S. Higashi, K. Raghavachari, and V. A. Burrows, *J. Vac. Sci. Technol. A* **7**, 2104 (1989).
- ¹³Y. J. Chabal and K. Raghavachari, *Phys. Rev. Lett.* **54**, 1055 (1985).
- ¹⁴The mode at 2074 cm^{-1} is the antisymmetric stretch ν_{as} of the coupled monohydride mode of steps on $\text{Si}(100)$; the symmetric mode ν_{ss} occurs at $\sim 2087\text{ cm}^{-1}$ [P. Jakob and Y. J. Chabal, *J. Chem. Phys.* **95**, 2897 (1991)] and cannot therefore be resolved from the tail of the 2091 cm^{-1} feature. The mode at 2098 cm^{-1} in the 650°C spectrum is the ν_{ss} of the coupled monohydride on the flat (100) terraces; the antisymmetric component is not observed due to the decreased relative sensitivity to such “parallel” modes upon developing micron-sized internal cracks (Ref. 10).
- ¹⁵G. S. Higashi, Y. J. Chabal, G. W. Trucks, and K. Raghavachari, *Appl. Phys. Lett.* **56**, 656 (1990).
- ¹⁶N. M. Johnson, F. A. Ponce, R. A. Street, and R. J. Nemanich, *Phys. Rev. B* **35**, 4166 (1987); H. P. Strunk, H. Cerva, and E. G. Mohr, *J. Electrochem. Soc.* **135**, 2876 (1989).
- ¹⁷P. Deák, L. C. Snyder, and J. W. Corbett, *Phys. Rev. B* **37**, 6887 (1988).
- ¹⁸A. J. Leadbetter, A. A. M. Rashid, N. Colenutt, A. F. Wright, and J. C. Wrights, *Solid State Commun.* **38**, 957 (1981).
- ¹⁹H. J. Stein, S. M. Myers, and D. M. Follstaedt, *J. Appl. Phys.* **73**, 2755 (1993).
- ²⁰Q.-Y. Tong, T.-H. Lee, K. Gutjahr, S. Hopfe, and U. Gösele (unpublished).

Minimization of suboxide transition regions at Si-SiO₂ interfaces by 900 °C rapid thermal annealing

G. Lucovsky,^{a)} A. Banerjee, B. Hinds, B. Claflin, K. Koh, and H. Yang

Departments of Physics, Materials Science and Engineering, Electrical and Computer Engineering, and Chemistry, North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 17 February 1997; accepted 10 April 1997)

Transitions regions at Si-SiO₂ interfaces contain excess suboxide bonding arrangements which contribute to interface roughness and also can give rise to electronically active defects. This article provides insights into the origin and temperature stability of these suboxide bonding arrangements by studying different interface formation processes, e.g., rapid thermal oxidation and plasma-assisted oxidation, and then subjecting these interfaces to rapid thermal annealing (RTA). The interfacial bonding chemistry has been studied before and after the RTA by Auger electron spectroscopy and it has been demonstrated that interfacial transition regions with suboxide bonding are a direct result of thermal and plasma-assisted oxidation at temperatures up to at least 800 °C, and that the excess suboxide bonding in interfacial transition regions is significantly reduced following a 30 s, 900 °C RTA. The kinetics of this interfacial annealing process are essentially the same as observed for the RTA-induced separation of homogeneous suboxide thin films (SiO_x, $x < 2$) into silicon nanocrystals and stoichiometric SiO₂. © 1997 American Vacuum Society.

[S0734-211X(97)07004-2]

I. INTRODUCTION

As channel lengths of field effect transistors (FETs) decrease into the deep submicron to 0.1–0.05 μm to provide increased levels of device integration, there must be accompanying decreases in gate oxide thicknesses (t_{ox}) initially to 3 nm and ultimately to <2 nm in order to maintain the drive currents needed for circuit operation. Once gate oxide thicknesses are reduced to a regime where direct tunneling occurs, e.g., <3 nm, the planarity of the gate oxide interfaces with both the channel region and gate electrode of the FET structure becomes a crucial issue since atomic scale variations in the effective gate oxide thickness can lead to order of magnitude variations in local tunneling currents. This article focuses on excess suboxide bonding arrangements that contribute to atomic scale roughness at Si-SiO₂ interfaces. It is important to recognize that interface roughness at the SiO₂-gate electrode interface must also be addressed since variations in the tunneling thickness can be induced by the roughness at both interfaces of an ultrathin gate oxide.

At an ideal Si(111)-SiO₂ interface with single Si dangling bond termination, the bonding arrangements can be characterized as Si-Si₄ in the Si substrate (the subscript indicates the number of Si atoms bonded to the reference Si atom), Si-O₄ in the bulk of the oxide, and Si-Si₃O at the metallurgical boundary. The Si-Si₃O bonding arrangement is one of three suboxide bonding arrangements; the other two are Si-Si₂O₂ and Si-SiO₃. If the interface between the Si substrate and the SiO₂ is not perfectly flat, then there must be additional or excess suboxide bonding arrangements that define a *transition region* between the Si substrate and a stoichiometric *bulk* oxide. This means that there is inherently a direct relationship between *excess* suboxide bonding ar-

rangements, interfacial transitions regions, and interfacial roughness. Assuming that the top surface of the oxide is perfectly flat, localized tunneling (either direct or Fowler-Nordheim) above an average value will take place at those interface locations where these excess suboxide bonding arrangements *extend into* the oxide and reduce the thickness through which electrons tunnel.

The issue of interfacial roughness and transition regions with excess suboxide bonding is further complicated by experimental observations that interfacial nitridation improves device reliability, at least in N-channel devices.^{1,2} This means that interface roughness may involve both Si-O and Si-N bonding arrangements in the nitrified structures. This article addresses roughness and suboxide bonding at Si-SiO₂ interfaces prepared by (i) remote plasma-assisted oxidation (RPAO) using O₂ and N₂O source gases,³ and (ii) rapid thermal oxidation (RTO) in O₂ and N₂O, and compares results of these studies with thermal oxidation in O₂.

Section II presents new experimental data obtained by on-line Auger electron spectroscopy (AES) that demonstrates changes in bonding at Si-SiO₂ interfaces after interfaces formed by RPAO and RTO at temperatures of 300 and 800 °C, respectively, and subjected to 30 s RTAs at 900 °C. The aspect of the present work that distinguishes it from past studies is the systematic control or processing conditions, in particular the separate and independent control of interface formation by RPAO and RTO at temperatures ≤800 °C and interface relaxation by RTA at 900 °C. Section III compares the results of this study with recent results of other studies that have addressed the thermal stability of bulk homogeneous suboxides,^{4,5} as well as x-ray photoelectron spectroscopy (XPS) studies that have directly addressed chemical bonding in the interfacial transition regions.^{6–8} Interfacial roughness and suboxide transition regions have also been

^{a)}Electronic mail: gerrylucovsky@ncsu.edu

studied by x-ray scattering,⁹ medium energy ion scattering (MEIS),¹⁰ and transmission electron microscopy (TEM) imaging.¹¹ This paper also includes experimental results from second-harmonic generation (SHG) studies that bear on the issue of suboxide bonding (see Ref. 3 and Refs. therein).

II. EXPERIMENTAL RESULTS

The new AES results presented in this article have been obtained using an ultrahigh vacuum (UHV) compatible multichamber system that has separate chambers for remote plasma processing (RPP), rapid thermal processing (RTP), and on-line AES. This system provides for (i) fabrication of Si-SiO₂ interface structures by either RPAO or RTO, and then (ii) structural and chemical relaxation of these interfaces by RTA. Two different oxygen/nitrogen atom source gases have been used: O₂ and N₂O.³ For both the RPAO and RTO processes, the surfaces of high resistivity $\sim 10 \Omega \text{ cm}$ p-type Si wafers were prepared by a two-stage process: high temperature ($\sim 900^\circ \text{C}$) thermal oxidation to grow a sacrificial oxide layer that is subsequently removed by rinsing in dilute HF. Wafers were then inserted into the plasma processing chamber of the system via a load lock sample introduction chamber. For the RPAO process, they were heated to 300°C at the chamber base pressure, $\sim 10^{-8}$ Torr, and then subjected to plasma-assisted oxidation. The RPAO process has been previously used as the first step of a two step process for the formation of gate oxides.¹² In addition to forming a superficially thin oxide, $\sim 0.5 \text{ nm}$, the RPAO process reduces levels of interface contamination, e.g., residual C and F concentrations are at the 10^{12} cm^{-2} range.² In the application of RPAO to gate oxide fabrication, $\sim 0.5\text{--}6 \text{ nm}$ of oxide is grown at 300°C by remote plasma excitation of either He/O₂ or He/N₂O mixtures (rf power of $\sim 30 \text{ W}$ at 13.56 MHz) and this is followed by remote plasma enhanced chemical deposition (RPECVD) of the remainder of the oxide film.^{2,12} Thicker films with t_{ox} to $\sim 1.5 \text{ nm}$, have also been grown directly by the RPAO process by extending the oxidation time.³ The oxidation process kinetics have been characterized by a power law relation; e.g., for growth in O₂, $t_{\text{ox}} \sim 0.7t^{0.28}$, where t is the time in minutes. The power law representation means that the oxidation is very fast initially when plasma-generated excited O species interact with Si atoms that are not bonded to other O atoms, but then the oxidation rate slows down significantly when the oxidation involves Si atoms that are bonded to O as well as other Si atoms. The RTO process was performed at 800°C with a source gas flow rate of $2,000 \text{ sccm}$.

Figures 1(a) and 1(b) indicate oxide growth in using remote plasma excited mixtures of He/O₂ or He/N₂O, respectively, as monitored by on line AES. The oxide thickness is determined from the ratio of the relative amplitudes of the Si-Si AES feature at $\sim 92 \text{ eV}$ and the Si-O feature at $\sim 76 \text{ eV}$.¹² Films prepared using the He/N₂O source gas show nitrogen atom incorporation localized in the immediate vicinity of the Si-SiO₂ interface, as evidenced by a N_{KLL} AES feature and confirmed by secondary ion mass spectrometry (SIMS) and optical SHG.^{2,3} There is an AES spectral

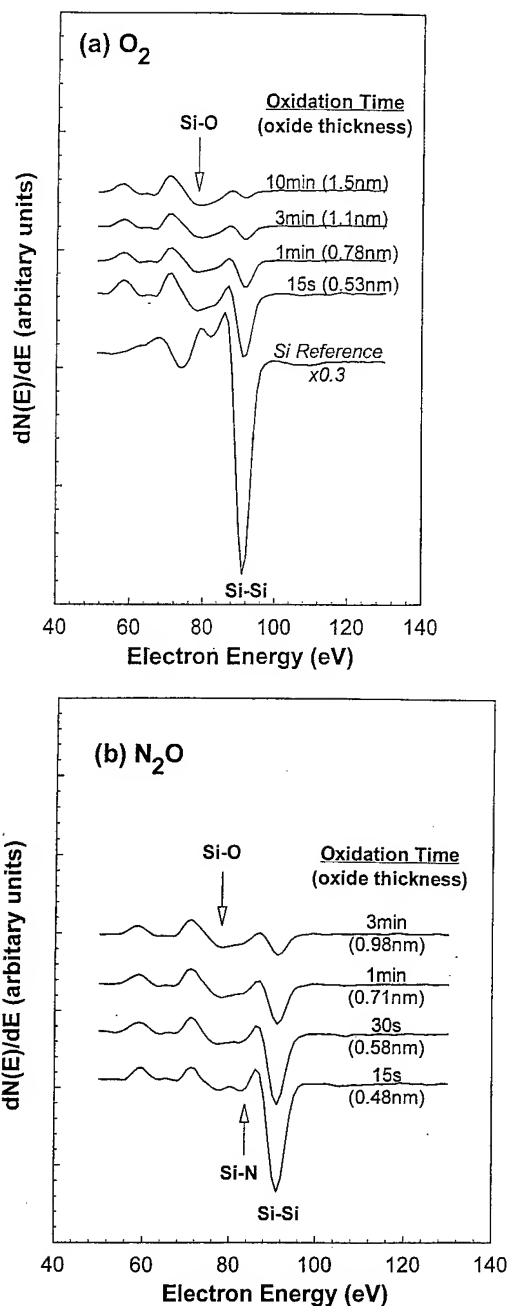


FIG. 1. Oxide growth in using remote plasma excited mixtures of (a) He/O₂ or (b) He/N₂O, respectively, as monitored by on line AES.

feature at $\sim 83 \text{ eV}$ in the Si_{LVV} region that has a contribution from Si-N bonding as well as an additional contribution from suboxide bonding arrangements. The contribution from suboxide bonding is identified by studies of the initial stages of oxidation before monolayer coverage has been obtained, whereas the contribution from Si-N has been verified by studying the interface between Si₃N₄ and Si. It is difficult to quantify the contribution to derivative AES spectra for the case in which both Si-N and suboxide bonding are present at the same time. However, as we show later on in the article, the relative changes in the derivative AES spectra after RTA are self-consistent, and are in agreement with the results of

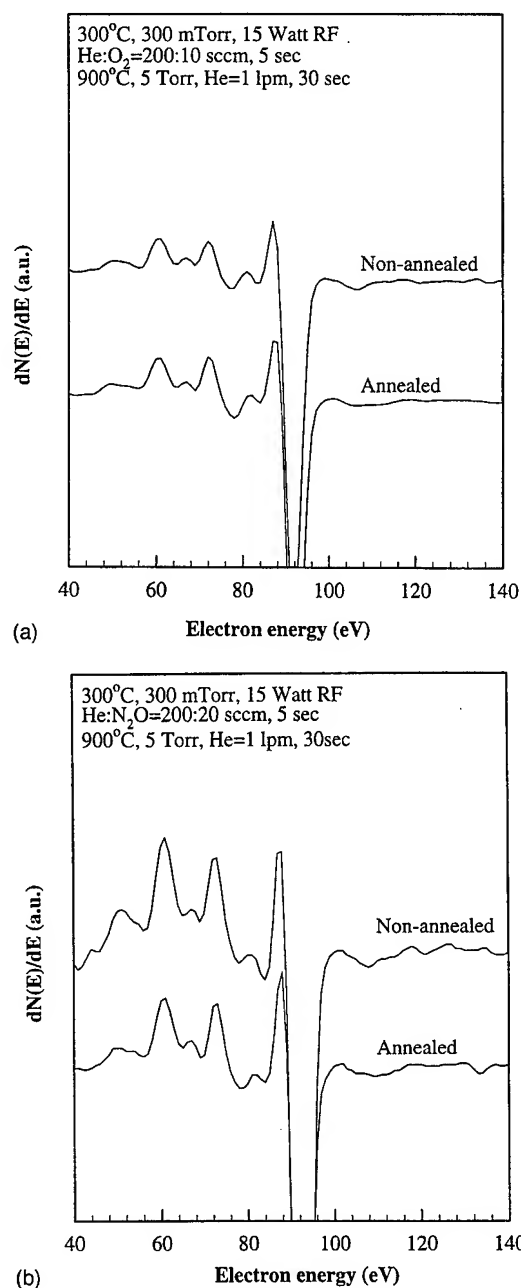


FIG. 2. Derivative Si_LVV AES spectra for oxide layers grown in (a) He/O₂ and (b) in He/N₂O, respectively, by RPAO that are ~0.5 nm thick as grown at 300 °C and after a 30 s 900 °C RTA.

other experimental probes of interface roughness before and after a 900 °C RTA or equivalent thermal exposure.

Figures 2(a) and 2(b) indicate derivative Si_LVV AES spectra for oxide layers prepared by RPAO using (a) He/O₂ and (b) in He/N₂O, respectively. These layers are ~0.5 nm thick as grown at 300 °C and after a 30 s 900 °C RTA. From the traces shown in Fig. 2 and the results in Table I, it is observed that the relative amplitude ratio, $[I(83 \text{ eV})/I(76 \text{ eV})]_{\text{after RTA}} / [I(83 \text{ eV})/I(76 \text{ eV})]_{\text{before RTA}}$, decreases after the anneal. Figures 3 and 3(b) indicate similar spectra for films prepared by RTO in O₂ and N₂O at 800 °C, and after a 30 s 900 °C RTA. Smaller decreases in the am-

TABLE I. Changes in the relative intensities of the SiO_x and SiO₂ AES at ~83 eV and 76 eV, respectively.

Interface formation	$\frac{[I(83 \text{ eV})/I(76 \text{ eV})]_{\text{after RTA}}}{[I(83 \text{ eV})/I(76 \text{ eV})]_{\text{before RTA}}}$
Plasma O ₂	0.88 ± 0.02
Plasma N ₂ O	0.79 ± 0.02
RTO O ₂	0.80 ± 0.02
RTO N ₂ O	0.86 ± 0.02

plitude ratio occur for these samples. To confirm that changes in the amplitude ratio were not due to oxidation during the RTA, the relative amplitudes of the Si-Si Si_LVV feature at ~92 eV, and the O_{KLL} feature at ~510 eV were also monitored and found to be the same before and after the RTA. Table I contains normalized amplitude ratios, $[I(83 \text{ eV})/I(76 \text{ eV})]_{\text{after RTA}} / [I(83 \text{ eV})/I(76 \text{ eV})]_{\text{before RTA}}$, for the results presented in Figs. 2(a), 2(b), 3(a), and 3(b). Since these ratios are all less than one, this means that there is a relative decrease in the amplitude of the 83 eV feature after the anneal.

III. DISCUSSION

Consider first the spectra in Figs. 2(a) and 2(b) for the as grown films. The feature at 76 eV is associated with Si-O bonding, in particular Si atoms with four oxygen atom neighbors. The AES results indicate that the feature at 83 eV has two contributions: (i) from Si-N bonds and (ii) from suboxide bonding arrangements, where the Si atom has one, two or three O atom neighbors. For example, AES results for films that display a relatively weak O_{KLL} feature, as in Fig. 2(a), also display a suboxide bonding feature at ~83 eV. In a more direct way, Tao *et al.*⁶ have shown significant reductions in suboxide bonding arrangements for films prepared by thermal oxidation at 600 °C, and then annealed at 925 °C consistent with changes in the AES features in films prepared by RTO and RPAO before and after 900 °C annealing. In particular, the concentration of suboxide bonding arrangements as determined from analysis of XPS data is reduced from ~3 monolayers by more than a factor of 2 after a 925 °C anneal. These changes in XPS are consistent with the AES intensity changes presented in Table I.

The effect of the 900 °C anneal in reducing suboxide bonding also helps to explain previously reported results from optical SHG measurements on oxidized vicinal Si(111) wafers off-cut in the 112 bar direction (Ref. 3 references therein). For the geometry of the measurements discussed in Ref. 3, the SHG signal comes predominantly from the Si-SiO₂ interface. These results are included in Table II which presents the results of experiments in which the azimuthal asymmetry has been addressed. The exciting source for these measurements is a Nd:YLF laser at 1.17 eV, and the Si(111) off-cut angles are approximately 5°. Experiments performed on thermally grown oxides subjected to rapid thermal annealing at 900 °C indicated a significant change in the rela-

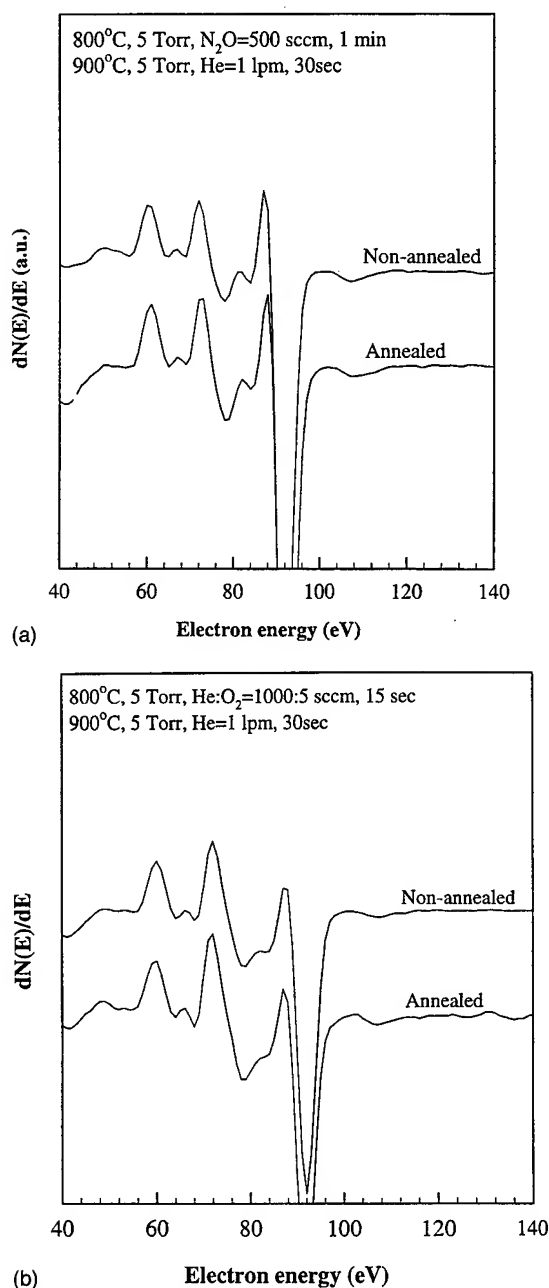


FIG. 3. Derivative Si_LVV AES spectra for oxide layers grown in (a) He/O₂ and (b) in He/N₂O, respectively, by RTO that are ~0.5 nm thick as grown at 300 °C and after a 30 s 900 °C RTA.

tive phase of the SHG signals associated with the steps and terrace regions.¹³ The SHG signal for the vicinal wafers is given by

$$E(2\omega) = (A_1 \cos \Phi + A_3 \cos 3\Phi) [\exp(-i\Delta_{13})], \quad (1)$$

where A_1 and A_3 are the amplitudes of the harmonic signal components at Φ and 3Φ , Φ is the angle between the 112 direction and the incident electric field, and Δ_{13} is the relative phase. The relative phase is related to the difference in resonance energies between the two harmonic components of the second harmonic signal. From Table II, the relative phase of the as-grown thermal oxide is 72°, and decreases signifi-

cantly to ~23° after the 900 °C anneal.¹³ In a similar way the relative phases for interfaces formed by 300 °C RPAO also change markedly after a similar 30 s 900 °C RTA. There are no significant differences between Δ_{13} for RPAO interfaces formed in N₂O and O₂, even though SIMS and on-line AES show a nitrogen terminated interface for the N₂O process. However, after the RTA, there are significant differences in Δ_{13} . Δ_{13} is equal to 23° for the oxygen terminated interfaces, the same value as obtained after annealing a thermally grown interface, but is reduced to 11° for the nitrogen terminated interface. The SHG results are then consistent with a significant change in interface bonding. Combining the SHG results with the XPS and AES results we conclude that the Δ_{13} values of 67°–72° are then characteristic of suboxide bonding arrangements, whereas the smaller values of Δ_{13} of 23° and 11° are indicative, respectively, of more nearly idealized interface bonding arrangements.

We have also studied the stability of suboxide films prepared by RPECVD at low temperatures (250–300 °C) and then subjected to RTA.^{4,5} The as-deposited films are hydrogenated suboxides. The local order in these films has been studied by infrared (IR) by analyzing the spectroscopic changes in the Si–H bond-stretching mode as a function of the oxygen concentration, x , as in SiO _{x} , $0 \leq x < 2$.^{14,15} This analysis indicated that the as-deposited films were random chemically-ordered alloys with statistically determined bonding of Si and O atoms in Si–Si₄, Si–Si₃O, Si–Si₂O₂, Si–SiO₃, and Si–O₄ local arrangements. Films for the annealing studies were prepared by RPECVD and then subjected to 30 s RTAs at 900 °C.^{4,5} Characterization of these films by IR indicated significant changes in the local bonding. Prior to the anneal, the films displayed a Si–O bond-stretching mode frequency between ~1010 and 1040 cm⁻¹ characteristic of homogeneous SiO _{x} alloys,^{14,15} and after the anneal the films displayed an Si–O bond-stretching mode frequency of ~1075 cm⁻¹, characteristic of stoichiometric SiO₂. Studies of the same films by TEM showed a homogeneous, noncrystalline character as-deposited, and a separation into Si nanocrystals (typical dimensions ~10 nm) embedded in a homogeneous amorphous matrix which was identified by IR as SiO₂.

The as-deposited homogeneous films showed strong photoluminescence at 80 K in a broad band extending from about 1.4 to 2.0 eV with the peak photon energy increasing as the x increased. This type of luminescence has been previously reported for homogeneous suboxide films.^{16,17} The annealed films displayed no detectable photoluminescence in the same spectral range. The absence of this type of luminescence in the annealed films is consistent with there being no significant suboxide regions at the interfaces between the Si nanocrystals and the encapsulating oxide films. Nanocrystals exposed to atmospheric oxygen and water vapor display characteristic suboxide bonding.¹⁸

Finally, recent studies of the kinetics of the process by which bulk suboxides, SiO _{x} , convert to Si nanocrystals encapsulated in SiO₂ or SiO _{x'} , ($x > x'$) has been studied.⁵ These studies have demonstrated that the separation process is lim-

TABLE II. Summary of results from studies of vicinal Si(111) surfaces off-cut $\sim 5^\circ$ in the 112 bar direction.

Surface treatment ^a	Phase, $\Delta\Phi^b$	A_1/A_3^c
Plasma processing		
O ₂ - 15 s - 300 °C	68	0.20 \pm 0.02
O ₂ - 15 s - 30 s 900 °C RTA (0.5O ₂ /Ar)*	23	0.35 \pm 0.03
N ₂ O - 15 s - 300 °C	67	0.21 \pm 0.02
N ₂ O - 15 s - 30 s 900 °C RTA (Ar)	11	0.37 \pm 0.04
N ₂ O - 30 s - 300 °C	65	0.17 \pm 0.02
N ₂ O - 30 s - 30 s 900 °C RTA (Ar)*	11	0.35 \pm 0.03
Thermally grown interfaces		
Furnace oxidation at 850 °C	72	0.19 \pm 0.02
Post-oxidation anneal [30 s at 950 °C (0.5%O ₂ /Ar)]	23	0.33 \pm 0.03

^aPre-deposition RPAO step.^bSee Eq. (1).^cProcessing conditions for optimum electrical properties (see Ref. 2).

ited by reaction kinetics, and that the time constants for the bulk separation at temperatures of the order of 900 °C are essentially the same as those for the interface relaxation process, e.g., complete separation takes place in ~ 30 at 900 °C.

There are three other recently reported studies that have addressed surface roughness/interfacial transition regions and changes that take place on annealing. Gibson *et al.* have used TEM interference techniques and demonstrated significant interfacial smoothing for thermally grown oxides subjected to a 900 °C anneal.¹¹ Downer and co-workers have used optical SHG to study Si-SiO₂ interfaces formed on Si(100) and have observed irreversible changes in the response from interfaces formed at low temperatures, ~ 800 °C, that were subsequently subjected to 900 °C thermal exposure.¹⁹ Finally, Matsumura and co-workers showed that ultrathin oxides (~ 3 nm thick) prepared by thermal oxidation at 650 °C and annealed at 850 °C showed improved performance with respect to stress-induced increases in tunneling currents.²⁰ They also found that oxides grown at 850 °C showed improved stress resistance after an anneal at 850 °C. These improvements were correlated with reductions in interface roughness as measured by x-ray scattering techniques.

IV. SUMMARY

The results presented in this article have demonstrated that for interfaces formed by conventional thermal oxidation, RTO and RPAO at temperatures below 800 °C show significantly reduced suboxide bonding after a relatively short RTA at 900 °C. Reductions in suboxide bonding, or equivalently increased interface smoothness, have been observed by a variety of different experimental techniques including (i) AES, (ii) XPS,^{6,7} (iii) optical SHG,³ (iv) x-ray scattering,⁹ and TEM.¹ Comparisons between the annealing of homogeneous bulk suboxide films, which results in the formation of Si nanocrystals encapsulated in SiO₂, and the interface relaxation process, suggest that both processes are limited by similar reaction kinetics.

The results presented in this article, combined with the results of other studies cited above indicate that the formation of suboxide bonding is inherent in the oxidation of crys-

talline Si independent of the oxidation process for oxidation processes that are performed at temperatures up to at least 900 °C. This has previously been reported for thermal oxidation, and the results in this article extend this observation to other oxidation mechanisms.^{21,22} If this is indeed the case, then the chemical and structural relaxations that take place during the 900 °C RTAs, or equivalent thermal exposures, will determine the limiting values of interface smoothness attainable. For example, suppose that oxidation takes place on an atomically smooth surface and that this process results in some degree of suboxide bonding. After the oxidation process is concluded there will be some degree of suboxide bonding/interface roughness. Following a 900 °C anneal, there will be interface smoothing accompanied by a reduction of suboxide bonding arrangements. However, since the roughening that takes place during oxidation, and the smoothing that takes place during the anneal are causing changes in atomic scale structure, it is more than likely that *complete* smoothing will not take place; i.e., that there will always be some degree of suboxide bonding above and beyond what is ideal for the particular crystallographic orientation of the Si substrate, independent of the initial state of the Si surface prior to oxidation. This explains an experimental observation that interfaces formed on smooth Si surfaces in different ways, e.g., by thermal oxidation, RTO, and RPAO, and then subjected to thermal exposures at 900–1000 °C after interface formation, show essentially the same degrees of interface roughness as monitored by the field dependence of the channel mobility in FETs.²³ Finally, the residual surface roughness and associated suboxide bonding arrangements may also be the origin of the limiting values of midgap interface defect states (D_{it}) of about 10^{10} cm⁻² eV⁻¹. The lack of distinct spectral feature in this midgap region suggests that the states contributing to D_{it} are not silicon atom dangling bond states. This has been confirmed by the recent experiments of Stathis and co-workers at IBM.²⁴ We suggest that the midgap D_{it} may be derived from changes in local dipoles that are induced by the applied electric fields as the bias voltage is swept from depletion to accumulation in the capacitance-voltage measurements. Clearly more work is necessary in this area.

This article has not addressed oxidation processes performed at temperatures in excess of 900 °C. For example, it would be interesting to determine whether high temperature thermal or rapid thermal oxidations at temperatures >950 °C induces roughness, i.e., excess suboxide bonding in interfacial transition regions, and to determine the extent to which a 900 °C RTA is effective in reducing any roughness produced.

ACKNOWLEDGMENTS

Support by the Office of Naval Research, the National Science Foundation, and the Semiconductor Research Corporation is greatly appreciated.

- ¹D. Mathiot, A. Straboni, E. Andre, and P. Debenest, *J. Appl. Phys.* **73**, 8215 (1993); J. Ahn, J. Kim, G. Q. Lo, and D.-L. Kwong, *Appl. Phys. Lett.* **60**, 2089 (1992).
- ²D. R. Lee, G. Lucovsky, M. R. Denker, and C. Magee, *J. Vac. Sci. Technol. A* **13**, 607 (1995); D. R. Lee, C. Parker, J. R. Hauser, and G. Lucovsky, *J. Vac. Sci. Technol. B* **13**, 1778 (1995).
- ³G. Lucovsky, H. Niimi, K. Koh, D. R. Lee, and Z. Jing, *The Physics and Chemistry of SiO₂ and the Si-SiO₂ Interface*, edited by H. Z. Massoud, E. H. Poindexter, and C. R. Helms (Electrochemical Society, Pennington, NJ, 1996), p. 441.
- ⁴A. Banerjee and G. Lucovsky, *MRS Symp. Proc.* **420**, 405 (1996).
- ⁵B. Hinds, G. Lucovsky, A. Banerjee, and R. Johnson, *MRS Symp. Proc.* (in press).
- ⁶H.-S. Tao, J. E. Rowe, H. Niimi, H. Yang, T. E. Madey, and G. Lucovsky (unpublished).

- ⁷F. J. Grunthaner and P. J. Grunthaner, *Mater. Sci. Rep.* **1**, 65 (1986).
- ⁸F. J. Himpsel, F. R. McFeely, A. Teleb-Ibrahimi, J. A. Yarnoff, and G. Hollinger, *Phys. Rev. B* **38**, 6084 (1988); M. M. Banaszak-Holl, S. Lee, and F. R. McFeely, *Appl. Phys. Lett.* **65**, 1097 (1994).
- ⁹M.-T. Tang, K. W. Evans-Lutterodt, G. S. Higashi, and T. Boone, *Appl. Phys. Lett.* **62**, 3144 (1993); M. L. Green, D. Brasen, K. W. Evans-Lutterholt, L. C. Feldman, K. Krisch, W. Lennard, H.-T. Tang, L. Manchanda, and M.-T. Tang, *ibid.* **65**, 848 (1994).
- ¹⁰E. P. Gusev, H. C. Lu, T. Gustafsson, and E. Garfunkel, *Phys. Rev. B* **52**, 1759 (1995).
- ¹¹X. Chen and J. M. Gibson, *Appl. Phys. Lett.* **70**, 1462 (1997).
- ¹²T. Yasuda, Y. Ma, S. Habermehl, and G. Lucovsky, *Appl. Phys. Lett.* **60**, 434 (1992).
- ¹³C. H. Bjorkman, T. Yasuda, C. E. Shearon, Jr., U. Emmerichs, C. Meyer, K. Leo, and H. Kurz, *J. Vac. Sci. Technol. B* **11**, 1521 (1993).
- ¹⁴S. S. Chao, J. E. Tyler, Y. Takagi, P. G. Pai, G. Lucovsky, S. Y. Lin, C. K. Wong, and M. J. Mantini, *J. Vac. Sci. Technol. A* **4**, 1574 (1986).
- ¹⁵D. V. Tsu, G. Lucovsky, and B. N. Davidson, *Phys. Rev. B* **40**, 1795 (1989).
- ¹⁶M. A. Paesler, D. A. Anderson, E. C. Freeman, G. Moddel, and W. Paul, *Phys. Rev. Lett.* **41**, 1492 (1978).
- ¹⁷J. C. Knights, R. A. Street, and G. Lucovsky, *J. Non-Cryst. Solids* **35-36**, 279 (1980).
- ¹⁸Y. Kanemitsu, H. Uto, Y. Masumoto, T. Masumoto, T. Futagi, and H. Mimura, *Phys. Rev. B* **48**, 2827 (1993).
- ¹⁹J. Didap, X. F. Hu, M. H. Anderson, M. C. Downer, M. terBeek, J. K. Lowell, and O. A. Aktsipetrov, in Ref. 2, p. 406.
- ²⁰T. Sakoda and Matsumura, *Proceedings of ISCSI*, Karuizawa, Japan, 1996.
- ²¹F. M. Ross, J. M. Gibson, and R. D. Twisten, *Surf. Sci.* **310**, 243 (1994).
- ²²P. O. Hahn and M. Henzler, *J. Vac. Sci. Technol. B* **2**, 574 (1984).
- ²³J. R. Hauser (private communication).
- ²⁴J. Stathis (unpublished).

Investigation of existing defects and defect generation in device-grade SiO₂ by ballistic electron emission spectroscopy

H. J. Wen and R. Ludeke

IBM T. J. Watson Research Center, Yorktown Heights, New York 10598

(Received 13 January 1997; accepted 3 April 1997)

Degradation processes initiated by defect generation in device-grade SiO₂ were studied by locally injecting hot electrons from a scanning tunneling microscope tip into Pd/SiO₂/p-Si(100) metal-oxide semiconductor (MOS) structures. An analysis of the emerging collector current in the Si substrate, a technique known as ballistic electron emission microscopy, provides electron transport information, from which the oxide defect generation process was studied. The charging of the defects resulted in shifts of threshold energies for electron transport across the oxide. A novel sheet charge model was developed to assess the in-depth distribution and charge densities in the oxide from field-induced threshold shifts obtained from experiment. An as-fabricated MOS system with an oxide thickness of 71 Å was investigated and found to contain existing electron traps of charge densities in the range $(0.7\text{--}2.8) \times 10^{13}$ e/cm² that are distributed within a 30 Å region adjacent to the metal/oxide interface. Further stressing was performed at zero oxide bias with increasing tip voltages of up to -10 V. New electron traps characterized by charge densities of $(1.9\text{--}3.6) \times 10^{13}$ e/cm² and located within 40 Å of the SiO₂/Si interface were generated when the kinetic energy of the electrons injected into the SiO₂ conduction band exceeded 1.9 eV. This energy threshold is in very good agreement with the hydrogen-release energy that is frequently invoked to explain oxide degradation. © 1997 American Vacuum Society. [S0734-211X(97)06704-8]

I. INTRODUCTION

The performance and reliability of modern metal-oxide semiconductor (MOS) devices are ultimately dependent on the electronic properties of the silicon dioxide, and those of its interfaces with the Si substrate and the gate metal.¹ The electronic breakdown characteristics, which are determined by defects and impurities in the SiO₂ bulk or at its interfaces, become increasingly sensitive to local fluctuations in material properties as device dimensions are reduced.²⁻⁵ The present understanding of degradation and breakdown mechanisms is largely based on studies on MOS capacitor structures with techniques such as Fowler-Nordheim (F-N) tunneling injection, internal photoemission, etc.⁶⁻²³ The results of these studies represent properties typically averaged over sampling areas of 0.01-1 mm², which do not necessarily reflect sample variations on a microscopic scale. Further shortcomings of the standard electron injection techniques, including F-N tunneling, require high fields across the oxide to energize the injected electrons. The high fields result in position dependent kinetic energy distributions of the electrons as they traverse across the oxide layer. For thicker oxides the energies may ultimately reach values sufficient to cause impact ionization in the oxide.⁶⁻¹⁰ Even at much lower fields the electrons still may reach energies sufficient to generate hot holes in the Si by impact ionization, which then can be injected into the silicon dioxide valence band.^{11,12} The presence of diverse types of hot carriers complicates the determination of oxide degradation mechanisms. The two most popular models, referred to as "impact ionization" and "trap creation," are both based on strong electron heating in the oxide.⁶⁻¹⁶ Impact ionization occurs when electrons in the oxide reach energies exceeding the SiO₂ band gap energy of

9 eV, which is necessary for creating electron-hole pairs in the oxide. Electron-hole recombination can then create interface states and electron traps. Defects in the trap creation model are generated by hot electrons breaking hydrogen-silicon bonds at the anode. The released hydrogen atoms diffuse to the cathode where they combine with other hydrogen atoms and escape, leaving defects behind. In this model the energy of hot electrons has a minimum threshold of ~2 eV in order to break H-Si bonds. However, under high field conditions the two defect generation mechanisms may coexist, which complicates the study of the many details of the processes. Moreover, such macroscopic results cannot precisely describe microscopic details of broken bonds, interface states, and charge traps. In order to understand oxide breakdown from a more fundamental point of view, studies on a microscopic scale become necessary.

In this article we present results from ballistic electron emission microscopy (BEEM) studies of both existing defects and the generation of new defects in gate oxides and near the interfaces of MOS structures. The results reported here were obtained for a 71 Å gate oxide. An "as-fabricated" MOS structure was found to have electron trap densities in the range of $(0.7\text{--}2.8) \times 10^{13}$ e/cm² that are distributed within a 30 Å region adjacent to the metal/oxide (M-O) interface. New electron traps, observed as trapped charge densities of $(1.9\text{--}3.6) \times 10^{13}$ e/cm² distributed within 40 Å of the SiO₂/Si interface, were generated when the energy of the hot electrons exceeded 1.9 eV. This value is close to the minimum energy of ~2 eV required to break hydrogen-silicon bonds.¹⁵ This process suggests that the initial defect generation, which eventually leads to a destructive breakdown of the oxide, is associated with the trap creation

model by releasing hydrogen. Hole injection and impact ionization are absolutely excluded in the present study, since the stressing was carried out at zero oxide bias and the kinetic energy of hot electrons was below the threshold for band gap ionization.

II. EXPERIMENTAL DETAILS

A. Ballistic electron emission microscopy/spectroscopy at various oxide fields

BEEM is a scanning tunneling microscopy (STM) based microscopy/spectroscopy. In this study electrons were injected from a STM tip into the metal layer of a MOS structure, as shown in Fig. 1(a). The sole purpose of this metal layer is to provide a potential reference plane, represented by the Fermi level in the metal, which is connected to ground in the present case. With the STM tip biased at V_T relative to the Fermi level in the metal, the electrons tunneling from the tip into the metal have a near-exponential energy distribution f_T with a maximum kinetic energy equal to eV_T , as shown in Fig. 1(b). In our BEEM configuration a maximum value of $V_T = -13$ V can be achieved, which corresponds to a kinetic energy of 9 eV relative to the conduction band minimum of the SiO₂ for the hottest injected electrons. For comparison, to reach this value in a 7 nm oxide by conventional F-N injection, a field of ~ 18 MV/cm would need to be applied. A separate gate bias V_b can also be readily applied between the metal and silicon substrate. This allows complete control over the field in the oxide and further modifies the electron energies. In general, the potential drop V_{ox} includes one part from the gate bias V_b and another part from the difference in workfunctions between the metal and Si substrate. Since V_b affects both V_{ox} and the band bending of the Si at the SiO₂-Si interface, the relationship between V_b and V_{ox} is nonlinear and can be obtained by solving Poisson's equation.^{1,24} Thus the energy of electrons injected into the SiO₂ conduction band cannot only be arbitrarily selected by changing the tip bias, but also independently modified by adjusting the oxide bias, thereby increasing the accessible energy range and allowing greater experimental flexibility.

The thickness of the metal layer should be of order of the electron scattering length in the metal, or preferably less, so that most of the electrons traverse the metal layer ballistically. In order to traverse the MOS structure and reach the Si substrate, the kinetic energy of the electrons must be larger than the potential barrier height in the MOS structure. The barrier height and shape in MOS structures is a much more complicated issue than for the metal/semiconductor system. It varies with the oxide bias, charge in the MOS system, work-function difference between the metal and the Si substrate, and image force effects at the interfaces. Figure 1(b) shows the schematic energy band diagram of the Pd/SiO₂/p-Si system at $V_b = 0$. By taking the potential step V_0 at the Pd/SiO₂ interface at zero oxide bias to be 4.08 eV,^{25,26} a band offset of 3.2 eV at the SiO₂/Si interface and a Fermi level location ~ 0.2 eV above the valence band maximum in the Si substrate, a nearly zero work-function difference between Pd and p-Si is deduced. Therefore, at

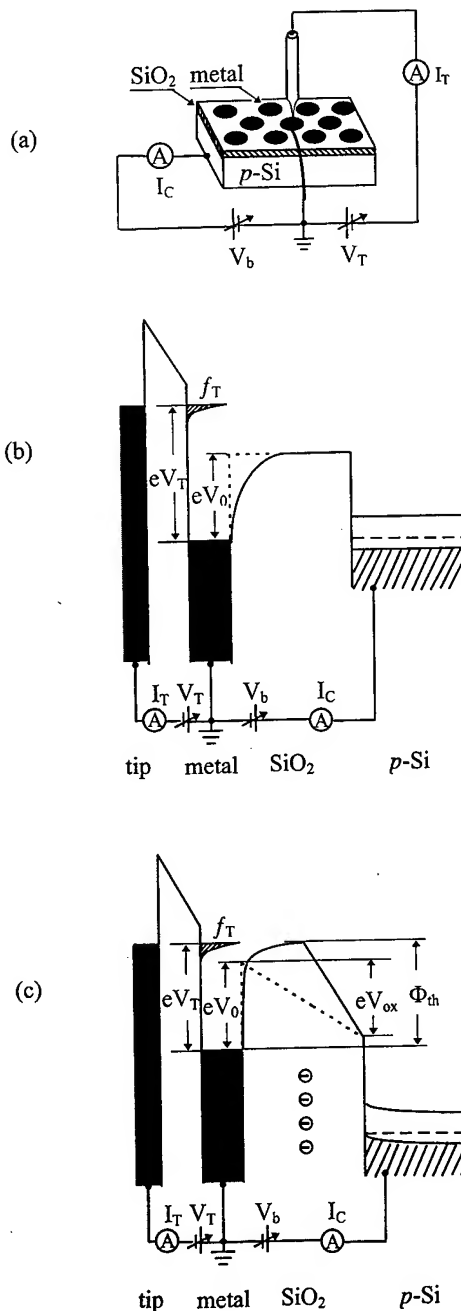


FIG. 1. (a) Schematic representation of BEEM experiment on a MOS structure, (b) schematic energy band diagram of a charge-free MOS system at $V_{ox}=0$, (c) energy band diagram of a MOS structure with a sheet of negative trapped charge in the oxide and under a positive oxide bias.

zero oxide bias the bands of the Pd/SiO₂/p-Si system are essentially flat. It should be noted that only image force effects, but no electron tunneling through the barrier^{25,26} are considered in Figs. 1(b) and 1(c). For the conditions of Fig. 1(b), electrons need a minimum kinetic energy of eV_0 to attain finite probabilities for crossing the M-O interface, traversing the oxide and being finally collected in the Si substrate as a collector current I_C . As a function of increasing V_T , I_C is zero until V_T reaches the value of the potential step V_0 ; thereafter it increases monotonically. The I_C vs V_T spec-

tra thus exhibit a threshold at V_0 . However, the location of the potential maximum and its height vary in the presence of an oxide charge. An energy diagram under positive oxide bias for an oxide film with negative trapped charge in the central region is shown in Fig. 1(c). One can readily realize that the effective barrier height Φ_{th} is strongly dependent on the charge density, charge distribution, and the applied oxide bias. Assuming the charge is distributed narrowly (sheet charge) and parallel relative to the interfaces, the effective barrier height lowering should then be linearly proportional to the applied oxide bias, with the slope of this relationship representing the location of the potential maximum as a fraction of the oxide thickness relative to the metal interface. In other words, charge density and charge location can be extracted from the field dependence of the effective barrier height. It should be noted that in Fig. 1(c) the sheet image charge in the metal is not depicted. Its inclusion is necessary for the interpretation of the BEEM results, and will be discussed in Sec. III. The experiment was devised as follows. As a first step, the as-fabricated MOS structures (without any hot electron stressing) were investigated. The existing defects manifest themselves as negative charge, which increases the effective barrier height and reduces the transmission probabilities. The location of existing defects was extracted by analyzing the field dependence of the barrier height. Then hot electrons were injected with a variable tip voltage of up to -10 eV in order to stress the oxide film. In this way, new defects were generated and further characterized by their field dependence.

B. Sample preparation

Device quality oxide layers were thermally grown near 800°C in dry oxygen on 125 mm diameter Si(100) wafers doped with boron to about 10^{16} cm^{-3} . No additional treatments were made after oxidation. Here we report studies on an oxide layer with a thickness $t_{ox} = 71 \pm 1 \text{ \AA}$. The thickness was determined with an ellipsometer. Approximately $2 \times 12 \text{ mm}^2$ pieces were cut from the wafers and introduced into an UHV chamber, where Pd dots, 0.2 mm in diameter, were thermally evaporated through a shadow mask. During deposition, the substrate was held at a temperature of about 30 K, which was necessary to inhibit surface diffusion. A peak-to-valley roughness of about 15 \AA was measured for a 40 \AA Pd film thickness. After evaporation, the sample was allowed to warm up to room temperature and was subsequently transferred under UHV into the STM chamber, where the BEEM grounding contact was gently positioned onto a selected Pd dot by means of three orthogonally mounted Inchworms. MOS structures with leakage resistance $\geq 10^{12} \Omega$ were generally suitable for our studies.

C. Data acquisition

All measurements were carried out when drifts caused by sample drift and piezocrystal creep were less than $1 \text{ \AA}/\text{min}$ in both the x and y directions. The low drifts were achieved either by waiting until the instrument was finally stabilized or by using drift compensation. In the spectroscopy mode the

raster scan of the STM is stopped and the collector current I_C is measured as V_T is ramped over a range that includes the barrier height. The barrier height manifests itself as a threshold voltage in the I_C vs V_T scan, from which the threshold is extracted with a model fit to the data.²⁴ The fitting process ensures negligible contributions to the threshold value from electrons tunneling through the barrier.^{25,26} The STM topographic images were taken in the conventional, feedback-controlled constant tunnel current I_T mode. I_C can also be recorded simultaneously, thereby generating an electron transmission or BEEM image of the MOS structure. Both point stressing and area stressing, achieved by injecting hot electrons at a point or over an area of $250 \times 250 \text{ \AA}^2$, were performed. In order to gain statistical information on a microscopic scale, I_C vs V_T curves were taken at 36 points of a grid pattern spread over a $250 \times 250 \text{ \AA}^2$ area of the surface, with a spacing of 40 \AA between points. It should be emphasized here that thresholds were obtained from single spectral scans (no averages), since we aimed to detect small changes that contain important information regarding the statistical nature of microscopic defects. Therefore all threshold energies presented here have an error bar of $\pm 0.05 \text{ eV}$. Further experimental details will be given as specific measurements are discussed. Details regarding the STM can be found elsewhere.²⁷

III. DATA INTERPRETATION AND DISCUSSIONS

A. Observation of existing traps and mechanisms of new trap generation

As mentioned earlier, the experiment was arranged to first investigate as-fabricated oxides before any intentional stressing. The existing traps are referred to as background or initial defects/impurities in the oxide and at the interfaces, which are incorporated into the structure during fabrication or processing and can be easily charged up by the injection of electrons or holes. These defects vary with the fabrication process and may be different in sign, density, and location from defects generated in aging studies. Therefore we intend to study first the existing, electrically active defects.

Figure 2 depicts a set of sequentially taken BEEM spectra that show a monotonic increase of the threshold with the number of spectral scan. These shifts clearly indicate that the mere spectrum acquisition process already begins to fill the pre-existing traps. During acquisition of these spectra the STM tip was located at the same point over the sample, the tip voltage was ramped only up to -6 V and the oxide bias was held at zero in order to prevent electron energy from exceeding 2 eV , which is considered a "soft" threshold for generating new defects. Each scan took less than 20 s. After the eighth scan the energy position of the threshold was stabilized. The increasing threshold indicates that the existing defects are electron traps. The final, stabilized threshold energy is $\sim 5.0 \text{ eV}$. This value, however, varies not only from one sample to another, as presented in Fig. 3 (the data in Fig. 2 were taken on sample 1 in Fig. 3) but may occur at different locations on a microscopic scale, as illustrated by the

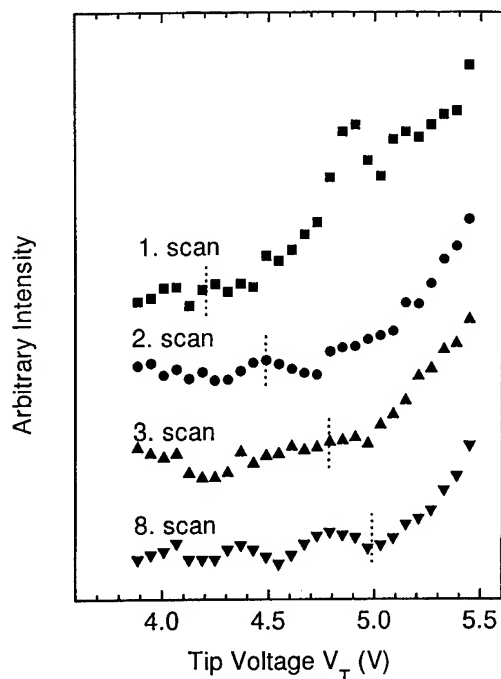


FIG. 2. BEEM spectra taken sequentially at the same sampling point of a Pd/SiO₂/p-Si structure. V_T ranged from -3 to -6 V and $I_T=2$ nA. It should be noted that only the range -3.9 – 5.5 V of V_T is shown here. Threshold energies of the spectra are marked by vertical dotted bars.

spread in thresholds depicted in Fig. 4. Such variations reflect a distribution in the local charge density of the existing traps, as will be presented in Sec. III D.

After the existing traps were saturated with electrons, i.e., no further shifts were observed in the BEEM spectra, hot electrons were injected with the STM tip on purpose to stress the oxide. The stress-induced defects manifest themselves as further shifts in the threshold energy, as shown in Fig. 3. The electron kinetic energy on the abscissa of Fig. 3 is given by $eV_T - e\Phi_{th}$. The zero threshold shift corresponds to the Pd–SiO₂ barrier height of 4.08 eV under flat band conditions.^{25,26} As discussed in Sec. II [Fig. 1(b)], the flat band condition for Pd/SiO₂/p-Si corresponds to $V_b=0$ V,

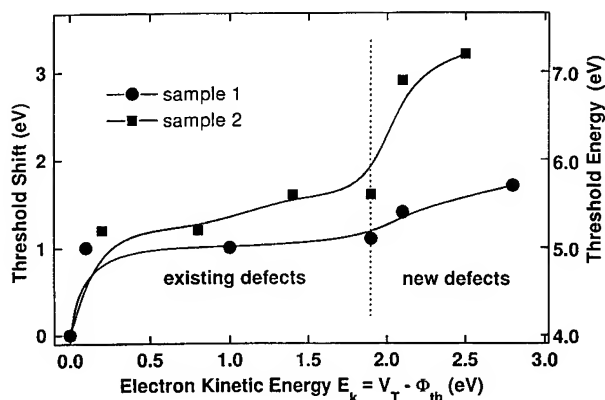


FIG. 3. Charge induced threshold shifts as a function of electron kinetic energy $E_k = V_T - \Phi_{th}$. Both curves corresponding to sample 1 (circles) and sample 2 (squares) show a soft threshold at ~ 1.9 eV beyond which the BEEM thresholds again increase.

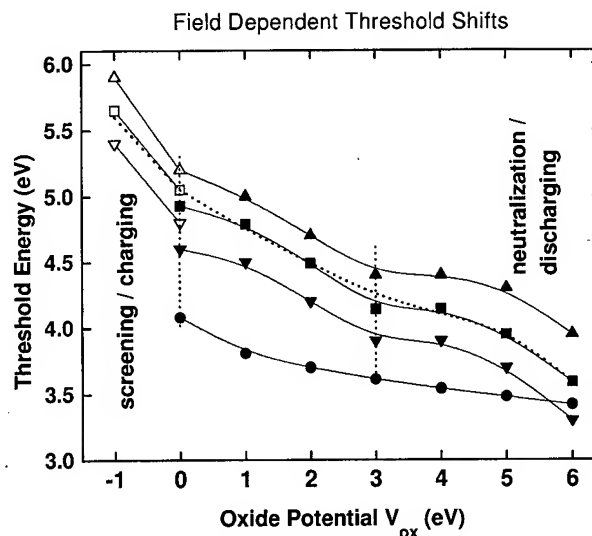


FIG. 4. Threshold shift as a function of applied oxide bias obtained from an area of 250×250 Å² where existing traps were filled up with electrons at $V_T = -6$ V. The area was sampled in a grid pattern of 36 spectra. The observed threshold shifts exhibited Gaussian distributions. The symbols correspond to the maximum (\blacktriangle) and minimum (\blacktriangledown) threshold shifts, as well as to the distribution maximum (\blacksquare) within a 36 spectral data set. The threshold shifts due to pure image force effects are also shown (\bullet). The dotted curve was obtained with the sheet charge model presented in Sec. III D. Two dotted vertical lines indicate three different regions of the field dependence.

with the assumption that no trap filling has occurred. Thus a deviation from the 4.08 threshold gives us an indication that the sample area below the tip is charged. For example, the first spectrum of Fig. 2 shows the threshold of 4.20 ± 0.05 eV, indicating a nearly charge-free area. The curves in Fig. 3 show a soft threshold of ~ 1.9 eV for the kinetic energy of the injected electrons. The curves continue to increase above this threshold, which indicates that more defects are generated by the increasingly hotter electrons. It must be noted that several spatially separated places were investigated during hot electron stressing. For a given electron kinetic energy above ~ 1.9 eV, fluctuations in the BEEM threshold with values as low as 5.0 eV were observed at several stressed points. Such extreme low values are attributed to leakage of the trapped charge, a process believed to occur in the prebreakdown region.²⁸ The data presented in Fig. 3 correspond to the highest value measured at several stressing places for a given electron kinetic energy. The soft threshold of ~ 1.9 eV can be directly related to the trap creation model. This model implies that traps are created by hot electron-initiated breaking of the hydrogen-silicon bonds at the SiO₂/Si interface, a process that needs an energy of ~ 2 eV.^{14–16,23} Consequently, we can postulate that the newly generated traps are located near the SiO₂/Si interface, a premise that is substantiated by the field dependent experiment presented in Sec. III C.

Since the kinetic energy for bond breaking is referenced to the potential maximum, electrons of energies exceeding 1.9 eV are hot enough anywhere in the oxide to release hydrogen. In contrast, for F–N injection, the field accelerated electrons have a higher probability to release hydrogen as

they move towards the anode. The defect generation process for the present case can be described as follows. Hot electrons with $E_k > 1.9$ eV release hydrogen both at the M–O interface and the SiO₂/Si interface. Since the M–O interface is believed to be a much hydrogen-rich source,²³ hydrogen atoms diffuse from there to the SiO₂/Si interface, where they combine with other released hydrogen and escape. Both the hydrogen releasing and recombination processes are potential defect generation mechanisms.

B. Location of existing traps

The trap densities and locations, as well as their polarity (electron or hole traps), are the relevant issues in trap characterization studies. In order to characterize existing traps, the following experiment was devised.

A pristine area of $250 \times 250 \text{ \AA}^2$ was first investigated in the spectroscopy mode by measuring a set of 36 spectra in a grid pattern. The tip voltage was ramped from -2 to -6 V in order to avoid new defect generation. It was observed that only a few, but always the first of the 36 spectra showed a threshold energy of 4.05 ± 0.05 eV, whereas most subsequent spectra already exhibited increases in thresholds of ~ 0.5 eV. This implies that the electrons injected into the SiO₂ conduction band during acquisition of the first spectrum quickly fill nearby existing traps and that this sequence is repeated during subsequent spectral scans. This process also indicates that in the present case the electron beam spread in the oxide exceeds the sampling spacing of 40 \AA , as was also observed in a previous study.²⁹ The area was then scanned with a tip voltage of -6 V until no further shifts were observed in the threshold energies, which indicates that saturation of the trap filling process was achieved. It must be noted that after saturation the spread in threshold energies of the 36 spectra for each data set exhibited an approximate Gaussian distribution. Their maximum, minimum threshold energies, and that at the distribution maximum are indicated in Fig. 4 by the up and down triangles and square symbols, respectively. The achievement of saturation means that the distribution does not change with further scans; it does not imply that each individual spectrum shows exactly the same threshold energy as the previous one, since the STM tip cannot exactly approach the same 36 sampling points between consecutive data sets. Such statistical results are better than a result from only one stressing point, since they provide a distribution of the existing trap densities over a microscopic area.

The oxide field dependent experiment was then performed in order to determine the trap distributions and densities. Identical grid patterns of 36 spectra were taken for gate biases in the sequence of 1 V steps from 0 through 7 V and back to 0 and -1 V, which corresponds to external oxide potentials of the same values. This simple relationship results from the fact that the Fermi level at the interface remains near the top of the Si valence band (in accumulation) due to the internal oxide field caused by the trapped charge. The threshold energies for $V_{ox} = 7$ V were difficult to determine due to a relatively large background leakage current and

therefore were not included in Fig. 4. The corresponding STM topographic images, generated concurrently with the spectral set, were used to check if the tip stayed at the same area. The results are presented in Fig. 4. It is seen that all three threshold data sets; minima, maxima, and distribution maxima of the 36 spectra exhibit almost the same field dependence and are, moreover, equally separated in energies. This well behaved and scatter-free behavior leads to the following conclusions: (a) The consistent spread between extrema and the distribution maxima, as well as their smooth dependence on V_{ox} , are well represented by the limited sampling size of 36, that is, the data is statistically meaningful. (b) Although each data set was not taken at the same points, the locations where the data sets were taken statistically exhibited very similar electrical environments. (c) No further trap filling takes place when V_{ox} increases from zero. (d) The distribution of the threshold energies represents a local fluctuation of trapped charge densities. In comparison, the pure image force lowering as a function of the oxide field is also indicated by the solid circles in Fig. 4. One readily sees that not only its field dependence differs from that of the three data sets, but that the shifts are also much smaller. It should be noted that this data set was taken from areas where no trapped charge was present, as detailed in our previous work.^{25,26,30}

As mentioned in Sec. II A, the magnitude of the field-dependent threshold shifts depends on the location of the potential maximum, which in turn depends on the charge distribution in the oxide. The field dependence of the three parallel data sets in Fig. 4 exhibits approximately three regions, as distinguished by vertical dotted lines. The region of V_{ox} ranging from 0 to 3 V has a nearly linear behavior with a slope of ~ 0.3 , while the region for $V_{ox} > 3$ V first shows a segment of weak field dependence that is followed by one of strong dependence. When the oxide potential V_{ox} changes from 0 to -1 V, the threshold energies increase by ~ 0.6 eV, a value that is twice the 0.3 eV shift expected solely from the applied oxide bias.

We can understand the overall field dependence through a simple model that will be detailed in Sec. III D. The model assumes that the trapped charge is distributed over separated sheets of charge parallel to the metal/oxide interface. When there is only one sheet charge present the potential maximum is located at the sheet charge position, whereas in the presence of several sheets of charge the shape of the total potential depends on their relative strength. Since charge in a MOS structure is spatially distributed, a detailed rendering of its effect on the potential can be reasonably approximated by considering several sheets of charge. A few general observations of the field dependence of the thresholds are worth stating here. The 0.3 slope indicates that the potential maximum in the bias range $0 \text{ V} < V_{ox} < 3 \text{ V}$ is located around $0.3 \times 71 \approx 20 \text{ \AA}$ from the M–O interface. This is a direct consequence of the reduced leverage of the applied oxide bias in affecting the potential maximum, whose position relative to the M–O interface is directly proportional to the slope of the field dependent threshold shifts. Consequently, the closer the

barrier maximum is to the M–O interface, the smaller will be the barrier lowering for a given increase in oxide bias. The data up to the $V_{ox}=3$ V indicates that the effective barrier lowering caused by a 3 V bias is 1 V, which compensates the original 1 V shift due to trap filling (approximately the shift between the solid circle and the square at $V_{ox}=0$ V). However, for $V_{ox}>3$ V, the softening of the curves implies that the field no longer affects the threshold energies. The curves bend down again upon further increasing the bias. This is attributed to a neutralization process and/or a discharge under high fields. It must be noted that after finishing the measurement at the high fields and before the bias was set to -1 V, three data sets at $V_{ox}=0$ V were taken again. These points are represented by open symbols. The square and down triangles are located 0.2 eV higher than the original positions, which can be assigned to a few new traps generated during the measurement at high fields. These new traps are believed to be similar to those discussed in Sec. III C. The dramatic increase in the thresholds for negative bias can only partially be attributed to the filling of additional electron traps close to the oxide/silicon interface. The additional charge required to fully account for the observed shift would exceed 10^{13} e/cm², a value that requires an extremely high density since the shift of the Fermi level in the Si at the interface is at most a small fraction of a volt by changing the bias from 0 to -1 V (the reader is reminded that the p -type Si at the interface is in accumulation). We attribute the excess in threshold shifts for negative bias as being predominantly caused by a change in the screening at the Si–SiO₂ interface from metallic to dielectric as the bias polarity is changed from positive to negative. This topic will be discussed in Sec. III D.

C. Location of generated traps

A procedure similar to that for the pre-existing defects was followed to characterize new traps generated by hot electrons of energies exceeding 1.9 eV. A pristine area of $250 \times 250 \text{ \AA}^2$ was again chosen and characterized spectroscopically in a grid pattern of 36 sampling points that exhibited a few sites with low thresholds of ~ 4.0 eV. Hot electrons were then injected over the entire area in sequences of increasing tip voltages of up to -10 V and with $V_b=0$ V. 36 spectra over the usual grid pattern were taken after each injection sequence in order to check threshold shifts. These followed a behavior similar to that shown in Fig. 3. Finally, the area was further scanned with $V_T=-10$ V until no more shifts in the threshold distributions were observed. For $V_{ox}=0$ the saturated distribution corresponds to the maximum and minimum threshold energies of 7.2 and 6.0 eV, as indicated by up and down triangles in Fig. 5, respectively. In comparison to Fig. 4, the larger threshold shifts clearly indicate that new electron traps were generated by the hot electrons. Then the field dependent experiment was performed. Spectra over the usual 36 sampling points were then taken in a series with V_{ox} equal to 0, 1, 2, 3, 4, 5, 6, 7, 0, and -1 V. Due to instabilities of the trapped charge, data sets for $V_{ox}>2$ V were so noisy that only a few of the 36 spectra

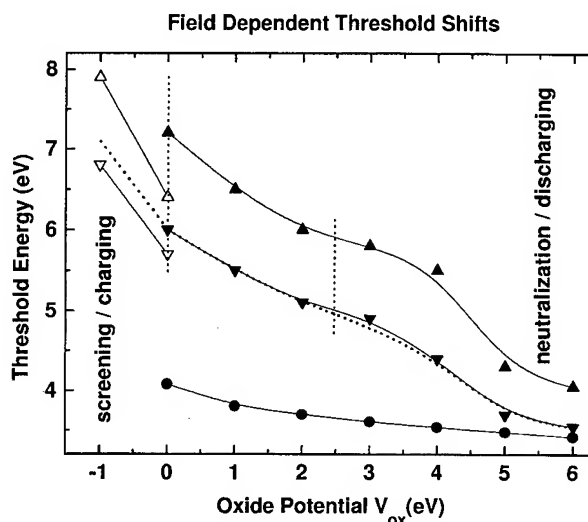


FIG. 5. Threshold shift as a function of applied oxide bias obtained from an area of $250 \times 250 \text{ \AA}^2$ that was stressed by hot electrons with V_T up to -10 V. The area was sampled in a grid pattern of 36 spectra. The threshold shifts exhibited Gaussian distributions. Up and down triangles and circles correspond to the maximum and minimum threshold shifts and threshold shifts due to pure image force effects, respectively. The dotted curve was obtained with the sheet charge model presented in Sec. III D. Two dotted vertical lines indicate three different regions of the field dependence.

exhibited reliable threshold characteristic. Consequently not enough data were available to determine a statistic distribution of thresholds. Therefore only the maximum and minimum threshold energies are presented in Fig. 5. The observed noise in the spectra is attributed to charging and discharging of the new traps at high oxide fields, a process that can happen when the trapped charge is located close to the oxide/silicon interface, or to charge neutralization via hot hole injection from the Si valence band. Although the data sets are not extracted from good statistical distributions, they still show the same field dependent behavior, which leads us to conclude that their separation adequately describes the maximum and minimum densities of the trapped charge. One sees in Fig. 5 that the field dependence is in general similar to that of Fig. 4, but the magnitude of the shifts is considerably larger. In comparison, the field dependent pure image force effect is also shown by the solid circles in Fig. 5. The observed field dependence is also divided into three regions, a nearly linear region with a slope of ~ 0.6 between 0 and 2.5 V, a region for $V_{ox}>2.5$ V that exhibits first a weak field dependence and then a much stronger one, and another region for V_{ox} between 0 and -1 V where the threshold energy dramatically increases by ~ 1.3 V.

From the slope of 0.6, we deduce that the potential maximum for V_{ox} between 0 and 2.5 V is located at $\sim 30 \text{ \AA}$ from the SiO₂/Si interface. A softening of the field dependence at ~ 2.5 V indicates that the maximum in the potential barrier shifts closer to the M–O interface, which implies a distributed oxide charge. At still higher oxide fields the increased rate of threshold shifts is again interpreted in terms of charge loss (neutralization or/and discharging). This high-field induced effect is also reflected in the ~ 0.5 eV lower threshold

energies at $V_{ox}=0$ V after high fields have been applied (open symbols). The dramatic increase for negative bias is largely attributed, as will be discussed in the next section, to changes in the screening mechanism at the SiO₂/Si interface.

D. Sheet charge model: charge density determination

Although oxide charge is expected to be distributed over some depth, it nevertheless can be reasonably approximated by one or more sheet charges so located that their resulting potential can describe the experimental observations. From previous discussions of the field dependence of the BEEM thresholds (Figs. 4 and 5) for $0 < V < 3$ V, it was concluded that a constant slope implies a specific position of the potential maximum relative to the M–O interface. In this case a single sheet charge located at the potential maximum would give a suitable potential profile. We model a single sheet charge in the oxide by considering it as a disk of radius R parallel to the oxide interfaces, of charge density σ_j , and located at a distance d_j from the M–O interface. The potential $\phi_j(d_j-x)$ along the axial direction x in an infinite dielectric medium of dielectric constant ϵ_{ox} is given by³¹

$$\phi_j(d_j-x) = \frac{\sigma_j}{2\epsilon} \{ [(d_j-x)^2 + R^2]^{1/2} - |d_j-x| \}, \quad (1)$$

where $\epsilon = \epsilon_0 \epsilon_{ox}$. ϵ_0 represents the permittivity of free space. The real potential $\Phi_j(x, d_j)$ must be zero at both interfaces because of metallic screening. These boundary conditions are satisfied by using the method of images,³¹ which is accomplished by adding the potentials of the image charges, including those from higher order terms (images of images, etc.):

$$\begin{aligned} \Phi_j(x, d_j) = & \phi_j(d_j-x) - \phi_j(d_j+x) - s[\phi_j(2t_{ox}-d_j-x) \\ & - \phi_j(2t_{ox}+d_j-x) - \phi_j(2t_{ox}-d_j+x) \\ & + \phi_j(2t_{ox}+d_j+x)] + \dots \end{aligned} \quad (2)$$

Here $s = (\epsilon_s - \epsilon_{ox}) / (\epsilon_s + \epsilon_{ox})$ is the effective screening factor at the O–S interface, with ϵ_s an effective dielectric constant of the semiconductor at the interface. For accumulation (or strong inversion) ϵ_s is assumed to become very large, so that $s \rightarrow 1$ (metallic screening). For $\epsilon_s = \epsilon_{Si} = 11.8$, $s = 0.5$. The profile of the resulting potential for a single disk shaped sheet charge is triangular, with the maximum (apex) at the location d_j of the sheet charge. In general, one can approximate the potential from a charge distributed along x by summing over a series of potentials arising from j individual sheet charges: $\Phi_T = \sum \Phi_j$. The barrier potential U is obtained by adding Φ_T to the sum of the barrier, field and image force potentials for the injected electron:

$$U = \Phi_T + V_0 - V_{ox}x/t_{ox} - q/(16\pi\epsilon x). \quad (3)$$

As discussed earlier in reference to Fig. 4, the potential maximum, judged from the slope in 0–3 V bias region, is $\sim 0.3t_{ox}$ or ~ 20 Å from the M–O interface. A single sheet charge was placed at this position (d_1) and its strength adjusted so that the total barrier potential U in Eq. (3) matches that of the experiment for $V_{ox}=0$. A single sheet charge,

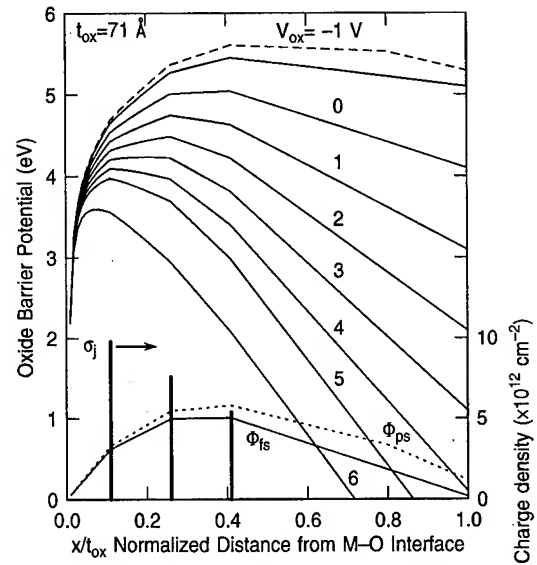


FIG. 6. Field-dependent barrier potentials calculated with the sheet charge model [Eq. (3)] for the saturated existing traps in a MOS structure with a 71 Å oxide. The heavy vertical bars show the sheet charge placements necessary to explain the experimental field-induced threshold shifts of Fig. 4. Their resulting potentials are shown at the bottom as Φ_{fs} and Φ_{ps} that represent, respectively, full and partial screening at the O–S interface. Partial (dielectric) screening applies only for the dashed curve for $V_{ox} = -1$ V; the solid curve assumes full screening.

however, overextends the region over which the slope equals 0.3 in Fig. 4. In order to account for the observed saturation for $V_{ox} > 3$ V, the potential Φ_T must bulge in the region of $x < d_1$. This occurs only if additional negative oxide charge is located within d_1 , which will be represented here by a second sheet charge placed at $d_2/t_{ox} = 0.11$ (~ 8 Å). Its density σ_2 is adjusted to mimic the softening observed in Fig. 4 for $V_{ox} > 3$ V. For a good general fit to the data, a third sheet charge was required beyond $d_1/t_{ox} = 0.3$, specifically at $d_3/t_{ox} = 0.41$. The σ_j 's were varied for an optimal fit to the data of Fig. 4 (solid squares), with relatively little freedom in the choice of the parameters. The net potential arising from the three sheet charges, indicated by the vertical bars in Fig. 6, is shown as curve Φ_{fs} , where the subscript refers to the choice of full (metallic) screening at both interfaces. The field dependent representations of the full barrier potential U [Eq. (3)] are shown as well in Fig. 6. The changing role of each sheet charge in determining the potential maximum as V_{ox} is increased from zero is clearly evident in the figure. The right-most sheet charge dominates for 0 and negative bias; the center charge dominates to $V_{ox} \sim 3$ V, beyond which the innermost sheet charge takes over. However, the subsequent rate of threshold lowering remains relatively weak and does not account for the drop beyond 5 V observed in Fig. 4. This can only be explained by a decrease in charge density (neutralization or discharging) under the highest field conditions ($V_{ox} = 6$ V), which can readily be simulated by lowering the net charge residing in the three sheets to 50% of its value. This reduction of the charge can conceptually occur either through field ionization or neutralization by hot holes injected into the oxide from the Si anode. Since pro-

TABLE I. Results from the sheet charge model for charge in an as-fabricated MOS structure, column 1, and charge generated by stressing, column 2. The space separates the model parameters of the pre-existing defects (top) from those of the generated defects. Three sheets of charge are used to represent the distributions of each type of defects. σ (range) is the range of total charge density necessary to represent the maxima and minima distributions for pre-existing (Fig. 4) and generated defects (Fig. 5).

	As-fabricated sample	Stressed sample
σ_j (e/cm ²) $\times 10^{12}$	9.8, 7.6, 5.4	9.8, 7.6, 5.4
d_j/t_{ox}	0.11, 0.26, 0.41	0.11, 0.26, 0.41
σ (e/cm ²) (range)	(7–28) $\times 10^{12}$	(7–28) $\times 10^{12}$
σ_j (e/cm ²) $\times 10^{12}$...	6.6, 6.9, 5.8
d_j/t_{ox}	...	0.55, 0.70, 0.85
σ (e/cm ²) (range)	...	(19–36) $\times 10^{12}$

longed high-field exposure without electron injection did not affect the thresholds measured subsequently at $V_{ox}=0$, we tend to ascribe the charge reduction to hole injection. Hole creation and injection is the result of electron-hole pair creation in the Si by impact ionization of hot electrons coming from the oxide. This process is quite efficient for electrons with kinetic energies >5 eV in the oxide.³²

As pointed out earlier in the discussions of Figs. 4 and 5, the large threshold shifts for negative bias appear to be anomalous, as the expected shifts would lie closer to values obtained by extrapolating the positive-biased segments of the curves near zero into the negative region. A sudden inrush of negative charge into the oxide at the O-S interface is unlikely, as this would suggest an interface trap density of the order of $10^{14}/\text{cm}^2$ eV near the Fermi level. However, the negative bias does have the tendency to move the Fermi level out of the valence band, which drastically affects the hole density in the Si at the O-S interface (because of the oxide charge the Fermi level remains near the top of the valence band). The decrease in the interface hole density has the effect of reducing the screening, which is described by the parameter s in Eq. (2). A value of $s=0.7$ was used to account for the observed threshold shifts for $V_{ox}=-1$ V. The resulting change in the oxide-charge potential is seen in the dotted profile marked Φ_{ps} in Fig. 6, as well as in the barrier potential shown by the top-most dashed curve. Φ_{ps} includes a contribution from a dipole charge of 3×10^{12} e/cm² located at $x/t_{ox}=0.8$ that is necessary to maintain band offsets at the interface. Its effect is seen by the slight kink in the two broken curves of Fig. 6. The maxima in the oxide barrier potentials of Fig. 6 are depicted in Fig. 4 by the dashed line. A similar treatment was followed for the generated defects, which were represented by an additional three sheet charges. We show here only the results, which are represented by a dashed line in Fig. 5 for the data corresponding to the lower limit in shift distributions. The parameters used to calculate the various potentials for both figures are summarized in Table I. It should be pointed out that other profiles and distributions, such as spreading the charge over much broader regions across the oxide, gave decidedly poorer fits to the experimental data, which in most cases

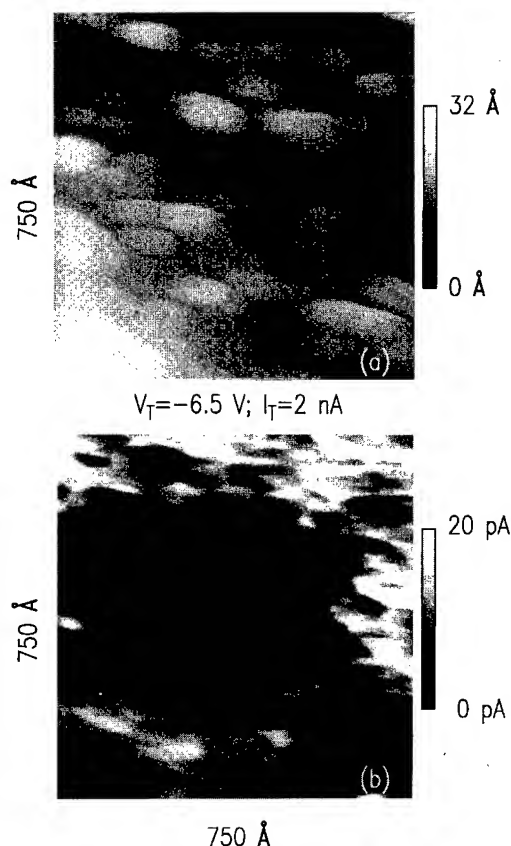


FIG. 7. (a) $750 \times 750 \text{ Å}^2$ STM topography and (b) BEEM image of a 40 Å Pd/ 71 Å SiO₂/p-Si structure stressed by hot electrons with V_T of up to -10 V. Black-white range over the image is 32 Å in topography, and 20 pA in BEEM. $V_T=-7$ V, $I_T=2$ nA, $V_b=0$ V. The stressed area is seen as a lower transmission region of $\sim 320 \text{ Å}$ diameter.

could only be fitted over a narrow bias range. The radius $R=16$ nm for all sheet charges was chosen to give an area equal to the scanned and stressed areas used in the experiments to generate the data for Figs. 4 and 5. This value was extracted from the transmission image of the stressed area shown in Fig. 7.

Several conclusions may be deduced from the results of the modeling. The obtained densities and location of existing traps near the M-O interface are consistent with previous observations on metal gated MOS structures.³³ Second, the trap densities near the M-O interface remain the same after the hot electron generation of new traps (in Table I, σ_j for existing traps remains the same in both columns). This means that hot electron stressing does not generate additional traps near the M-O interface, which supports our original interpretation that these traps pre-existed the BEEM experiment. Finally, the densities and location of the generated traps are not inconsistent with conclusions obtained with the hydrogen release model, which proposes a dangling bond origin of the traps as the hydrogen atom is stripped from the passivated Si-H bond near the SiO₂/Si interface.^{15,16}

E. Imaging charged traps

After completing the field dependent stress experiment, a large area of $750 \times 750 \text{ Å}^2$ that included the stressed area of

250×250 Å² was imaged simultaneously in the topographic and BEEM transmission modes, as shown in Fig. 7. The stressed area is seen as a low transmission region of approximately 320×320 Å². The observation of a low-transmission region that is larger than the stressed area is attributed to a lateral spread of the electron beam injected into the SiO₂ conduction band during the stressing and/or the data acquisition process.²⁸ It should be noted that before stressing the BEEM image of this area only exhibits a contrast resulting from the nodular character of the Pd surface, as is still evident in the periphery of the BEEM image. Stressing at one point was performed several times. An analysis of the threshold shifts in the proximity of the stressing point indicates a beam spread of ~100 Å, a value that can explain the large area of low BEEM transmission in Fig. 7. We would like to emphasize that at least for thick oxides, besides the influence of the metal topography on the BEEM image, such a big beam spread is another critical obstacle to the observation of individual traps, which would be a highly desirable achievement.

IV. SUMMARY

We have used a STM tip to locally inject electrons into MOS structures consisting of a 35 Å Pd layer on a 71 Å oxide, with the intent of studying defects and defect generation in SiO₂ gate oxides on a microscopic scale. Electron transport processes were studied by analyzing the BEEM collector current emerging from the Si substrate. In particular the shifts in thresholds were investigated that resulted from increases in the oxide potential due to the presence of trapped charge. A model was developed for the dynamic oxide potential profile that includes contributions from the oxide charge, field effects and image force (screening) effects from the interfaces. From the observed field dependence of the thresholds and suitable modeling, charge profiles and charge densities were deduced. The as-fabricated MOS structures exhibited charged electron trap densities of $(0.7\text{--}2.8) \times 10^{13}$ e/cm² that were distributed with decreasing density within 30 Å of the M–O interface. Stressing with hot electrons of energies exceeding 1.9 eV generated new electron traps, which manifest themselves as a trapped charge density of $(1.9\text{--}3.6) \times 10^{13}$ e/cm² located within 40 Å of the SiO₂/Si interface. The value of 1.9 eV is in excellent agreement with the minimum energy of ~2 eV required to break hydrogen-silicon bonds. The electrically stressed area was observed microscopically as a region of low transmission that results from the presence of electrons captured by the traps in the oxide.

ACKNOWLEDGMENTS

The authors benefited from discussions with D. J. DiMaria, D. A. Buchanan, and E. Cartier. This research was sup-

ported by the Office of Naval Research under Contract No. N00014-95-C-0087.

- ¹E. H. Nicollian and J. R. Brews, *MOS Physics and Technology* (Wiley, New York, 1982).
- ²C. Hu, *IEEE* **81**, 682 (1993).
- ³C. A. Mead, *Analog Integrated Circuits Signal Process.* **6**, 9 (1994).
- ⁴Y. Taur, Y.-J. Mii, D. J. Frank, H.-S. Wong, D. A. Buchanan, S. J. Wind, S. A. Rishton, G. A. Sai-Halasaz, and E. J. Nowak, *IBM J. Res. Dev.* **39**, 245 (1995).
- ⁵Semiconductor Technology Workshop, Working Group Reports, Semiconductor Industry Association, San Jose, 1993.
- ⁶P. Solomon, *J. Vac. Sci. Technol.* **14**, 1122 (1977).
- ⁷D. J. DiMaria, T. N. Theis, J. R. Kirtley, F. L. Pesavento, D. W. Dong, and S. D. Brorson, *J. Appl. Phys.* **57**, 1214 (1985).
- ⁸M. V. Fischetti, D. J. DiMaria, S. D. Brorson, T. N. Theis, and J. R. Kirtley, *Phys. Rev. B* **31**, 8124 (1985).
- ⁹D. Arnold, E. Cartier, and D. J. DiMaria, *Phys. Rev. B* **45**, 1477 (1992).
- ¹⁰D. J. DiMaria, D. Arnold, and E. Cartier, *Appl. Phys. Lett.* **60**, 2118 (1992).
- ¹¹K. F. Schuegraf and C. Hu, *IEEE Trans. Electron Devices* **41**, 761 (1994).
- ¹²K. F. Schuegraf, D. Park, and C. Hu, *Proceedings of the IEEE International Electron Device Meeting, 1994, IEDM 94*, p. 609.
- ¹³D. Arnold, E. Cartier, and D. J. DiMaria, *Phys. Rev. B* **49**, 10 278 (1994).
- ¹⁴E. Harari, *Appl. Phys. Lett.* **30**, 601 (1977).
- ¹⁵E. Cartier and D. J. DiMaria, *Microelectron. Eng.* **22**, 207 (1993); D. J. DiMaria, E. Cartier, and D. Arnold, *J. Appl. Phys.* **73**, 3367 (1993).
- ¹⁶D. J. DiMaria and J. W. Stasiak, *J. Appl. Phys.* **65**, 2342 (1989).
- ¹⁷P. Olive, T. N. Nguyen, and B. Riccò, *IEEE Trans. Electron Devices* **35**, 2259 (1988).
- ¹⁸P. Olive, T. N. Nguyen, and B. Riccò, *IEEE Trans. Electron Devices* **38**, 527 (1991).
- ¹⁹P. P. Apte, T. Kubota, and K. C. Saraswat, *J. Electrochem. Soc.* **140**, 770 (1993).
- ²⁰L. K. Han, M. Bhat, D. Wristers, J. Fulford, and D. L. Kwong, *Proceedings of the IEEE International Electron Device Meeting, 1994, IEDM 94*, p. 617.
- ²¹S.-H. Lee, B.-J. Cho, J.-C. Kim, and S.-H. Choi, *Proceedings of the IEEE International Electron Device Meeting, 1994, IEDM 94*, p. 605.
- ²²C.-H. Lin, J. Cable, and J. C. S. Woo, *IEEE Trans. Electron Devices* **42**, 1329 (1995).
- ²³D. A. Buchanan, A. D. Marwick, D. J. DiMaria, and L. Dori, *J. Appl. Phys.* **76**, 3595 (1994).
- ²⁴R. Ludeke, A. Bauer, and E. Cartier, *Appl. Phys. Lett.* **66**, 730 (1995); *J. Vac. Sci. Technol. B* **13**, 1830 (1995).
- ²⁵H. J. Wen, R. Ludeke, D. M. Newns, and S. X. Lo, *J. Vac. Sci. Technol. A* **15**, 784 (1997).
- ²⁶H. J. Wen, D. M. Newns, and R. Ludeke, *Phys. Rev. Lett.* (submitted).
- ²⁷M. Prietsch, A. Samsavar, and R. Ludeke, *Phys. Rev. B* **43**, 11 850 (1991).
- ²⁸R. Ludeke, H. J. Wen, and E. Cartier, *J. Vac. Sci. Technol. B* **14**, 2855 (1996).
- ²⁹B. Kaczer and J. P. Pelz, *J. Vac. Sci. Technol. B* **14**, 2864 (1996).
- ³⁰R. Ludeke, E. Cartier, and H. J. Wen, *J. Electrochem. Soc.* **96-1**, 580 (1996).
- ³¹R. H. Good and T. J. Nelson, *Classical Theory of Electric and Magnetic Fields* (Academic, New York, 1971).
- ³²D. J. DiMaria, E. Cartier, and D. A. Buchanan, *J. Appl. Phys.* **80**, 304 (1996).
- ³³D. J. DiMaria (private communications).

Macroscopic and microscopic studies of electrical properties of very thin silicon dioxide subject to electrical stress

E. S. Daniel, J. T. Jones, O. J. Marsh, and T. C. McGill^{a)}

T. J. Watson Sr. Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125

(Received 13 January 1997; accepted 17 April 1997)

The electrical characteristics of various size tunnel switch diode devices, composed of $\text{Al}/\text{SiO}_2/n\text{-Si}/p^+\text{-Si}$ layers, which operate with a range of parameters (such as current densities in excess of 10^4 A/cm^2) that stress the oxide layer far beyond the levels used in typical thin oxide metal-oxide semiconductor research have been examined. It is found that the first time a large current and electric field are applied to the device, a "forming" process enhances transport through the oxide in the vicinity of the edges of the gate electrode, but the oxide still retains its integrity as a tunnel barrier. The device operation is relatively stable to stresses of greater than 10^7 C/cm^2 areally averaged, time-integrated charge injection. Duplication and characterization of these modified oxide tunneling properties was attempted using scanning tunneling microscopy (STM) to stress and probe the oxide. Electrical stressing with the STM tip creates regions of reduced conductivity, possibly resulting from trapped charge in the oxide. Lateral variations in the conductivity of the unstressed oxide over regions roughly 20–50 nm across were also found.

© 1997 American Vacuum Society. [S0734-211X(97)06904-7]

I. INTRODUCTION

The continuing drive to decrease the characteristic dimensions of metal-oxide semiconductor-field effect transistors (MOSFETs) in modern integrated circuits has fueled basic research of metal-oxide semiconductor (MOS) structures with very thin ($<3.5 \text{ nm}$) silicon dioxide layers. Studies of current-voltage (I - V) and capacitance voltage (C - V) characteristics,^{1–7} as well as more basic surface studies using x-ray photoemission spectroscopy (XPS)^{4,8,9} and ballistic electron emission microscopy (BEEM),^{10–12} have greatly enhanced our understanding of current transport through these oxides, the chemical nature of the Si-SiO_2 interface, and the effects of electrical stressing. However, as the bulk of this work has been aimed at studying the future of the MOSFET, few studies have attempted to stress thin oxides ($<3.5 \text{ nm}$) with very high current densities ($>4 \times 10^3 \text{ A/cm}^2$). While an oxide that has been degraded with electrical stress is no longer usable in a MOSFET, it has not become an electrical short,^{2,4} and may be useful in other devices. Studies of such devices may prove useful in understanding the intrinsic properties of the oxide and its role in the operation of any device.

Consider the tunnel switch diode (TSD) device discovered by Yamamoto *et al.*¹³ [Fig. 1(a)], which has been proposed as a candidate for power switching, logic, and memory applications.^{14–16} The device consists of a p - n junction in series with a MOS junction, yielding a thyristorlike I - V characteristic shown in Fig. 1(b), consisting of two stable current states for a range of voltages. It has been demonstrated theoretically¹⁵ and experimentally^{17,18} that this device relies critically on the oxide as a tunnel barrier. If the oxide is either too thin or too thick, the peak in the I - V curve disappears, and there are no bistable current states for any

range in voltage. A comparison (Table I) of typical operating parameters of TSD devices produced in our laboratory with thin oxide MOS research parameters³ indicates that the TSD clearly operates in a parameter space far removed from that of the MOSFET, stressing the oxide far beyond the point at which the MOSFET fails.

In this article we present electrical studies of the TSD device, focusing on the role of the oxide and its ability to perform a critical role under heavy stress. By measuring the I - V characteristics of various size devices both before and after stress has been applied to the device, we find that initial stress causes a "forming" process, leading to nonuniform current transport through the device, but that further stress produces little or no change. This suggests a local modification of the oxide tunneling properties. In an attempt to create and characterize similar locally stressed regions in a somewhat simpler system, we examine a simple SiO_2 - Si interface using atomic force microscopy (AFM) and scanning tunneling microscopy (STM).

II. TSD STUDIES

A. Basic device operation

While the behavior of the TSD device is not completely understood, the generally accepted basic mechanisms leading to the thyristorlike I - V curve are described in detail elsewhere.^{14,15,18,19} Here we will only give a brief description of the device operation and the role of the thin oxide barrier.

The device is most easily described as a p - n junction and an MOS junction in series. If positive bias is applied to the gate electrode, the I - V curve is dominated by the reverse biased p - n junction and contains no unusual features. As negative bias is applied to the gate electrode, the MOS junction depletion layer begins to extend into the n - Si layer [Fig.

^{a)}Electronic mail: tcm@ssdp.caltech.edu

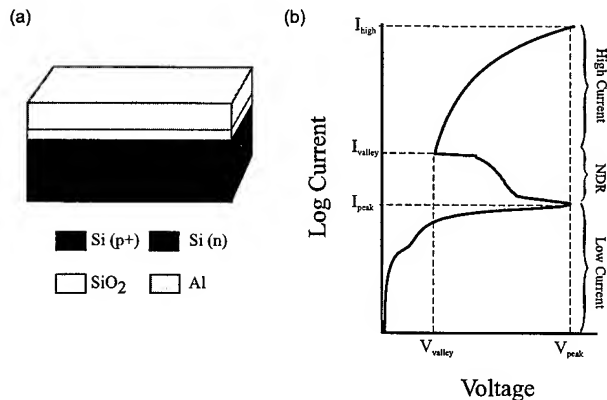


FIG. 1. Tunnel switch diode (TSD). (a) The TSD device consists of a p + silicon substrate with an n silicon epilayer, typically a few microns thick, a tunnel oxide (10–40 Å), and a conductive (e.g., Al) gate on top. (b) A typical measured I - V curve for a $150 \times 150 \mu\text{m}^2$ device. For this particular device, $I_{\text{high}} = 100 \text{ mA}$, $I_{\text{valley}} = 0.1 \text{ mA}$, $I_{\text{peak}} = 0.009 \text{ mA}$, $V_{\text{valley}} = 1 \text{ V}$, and $V_{\text{peak}} = 4 \text{ V}$.

2(a)]. In this state, the current is low, limited by the reverse-biased depletion layer, and the bulk of the bias is dropped across this layer. Once the depletion layer reaches the p - n junction depletion region, however, further bias causes the p - n junction to be turned on, flooding the n -Si layer with holes. If the oxide layer is not too thin, the holes will accumulate in the n region, creating an inversion layer. This shifts the bulk of the bias from the depletion layer to the oxide [Fig. 2(b)], causing a large increase in tunnel current, provided the oxide layer is not too thick. This, in turn, causes the Fermi level in the n -Si layer to be pulled up, forward biasing the p - n junction even more. This positive feedback mechanism switches the device into the high current state, in which the n -Si layer is nearly charge neutral, highly populated by both holes and electrons, with a hole inversion layer at the surface providing the large field across the oxide. This state can only be supported as long as the oxide is thin enough to allow large tunnel currents, but thick enough to support the hole inversion layer.

Other mechanisms can be employed to switch the device from the low to high current states. For example, light injection into the n -Si layer causes electron-hole pairs to be produced, setting up the necessary carrier distribution in the n -Si layer at a much lower bias.¹³ In Fig. 2(c), we see that increased light injection intensity results in a decrease in the

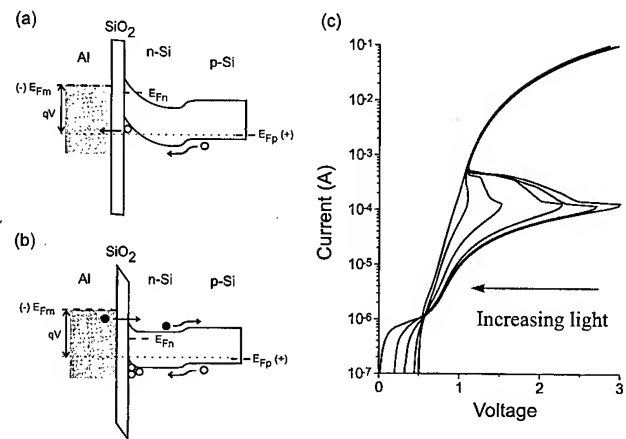


FIG. 2. (a) A band diagram of the TSD in the low current state just below the peak voltage. Further bias will extend the MOS junction depletion layer, forward biasing the p - n junction. E_{Fm} , E_{Fn} , E_{Fp} and V are the metal Fermi level, electron quasi-Fermi level at the SiO_2 / n -Si interface, the Fermi level in the flat band region of the p -Si substrate, and the applied bias, respectively. (b) A band diagram of the TSD in the high current state. The hole inversion layer maintains the majority of the applied bias across the oxide, while electrons tunneling through the oxide and holes injected from the p -Si layer keep the n -Si layer highly populated, but nearly charge neutral. (c) Measured I - V curves of a $150 \times 150 \mu\text{m}^2$ device for different intensities of injected light.

peak voltage of the TSD, until eventually the peak disappears completely. Charge injection via a third terminal, connected to the n -Si layer, results in a similar modification of the I - V characteristic.^{14,19}

B. Experimental procedures

The details of the device fabrication and measurement are described elsewhere¹⁷, so we will only reproduce a summary here. Chips approximately $1 \times 1.5 \text{ cm}^2$ were cleaved from commercially obtained wafers consisting of a p -type ($3 \times 10^{18}/\text{cm}^3$) silicon (100) substrate and an n -type ($2 \times 10^{15}/\text{cm}^3$) epilayer approximately $1.5 \mu\text{m}$ thick. After cleaning, thin sacrificial oxide layers were grown on the chips in a rapid thermal processing (RTP) oven in O_2 for 60 s at 900°C . During ramp-up and anneal in the oven, the chips were held in an Ar ambient. The sacrificial layers were stripped off in a buffered oxide etch (BOE), and another oxide was grown, again at 900°C in O_2 . The thickness of this oxide was controlled by varying the O_2 flow time. Samples with a flow time of 15, 30, 45, and 60 s were produced. The oxide thicknesses of these samples were estimated to be 1.5, 1.8, 2.1, and 2.5 nm, respectively, by measuring the thickness of the 60 s sample using standard C - V techniques and extrapolating along published calibration curves^{20,21} to the smaller thicknesses. As these values are only estimated and not measured, we will refer to the samples by the controlled O_2 flow time rather than the thickness. We note that the thickness estimation¹⁷ is calibrated using C - V measurements which have been reported to yield values roughly 0.9 nm less than those given by optical techniques such as ellipsometry when metal gates are used.²² Aluminum contacts were deposited on both the front and

TABLE I. A comparison of maximum stress operating parameters of the TSD devices produced in our labs with those of current thin oxide MOS research structures (see Refs. 2 and 5). For the purposes of a rough comparison, we define “ Q_{max} ” as the areally-averaged maximum fluence of charge that results in a nonfunctioning device, and we compare this with Q_{bd} in MOS devices. As our time-tested TSD devices did not expire in the course of our testing (Fig. 3), we can only give a lower bound on this number.

	t_{oxide} (Å)	Q_{max}, Q_{bd} (C/cm^2)	J_{oxide} (A/cm^2)	Electric field (MV/cm)
TSD	<35	$>10^7$	$>10^4$	10–20
MOS	30–60	$<10^5$	<10	<15

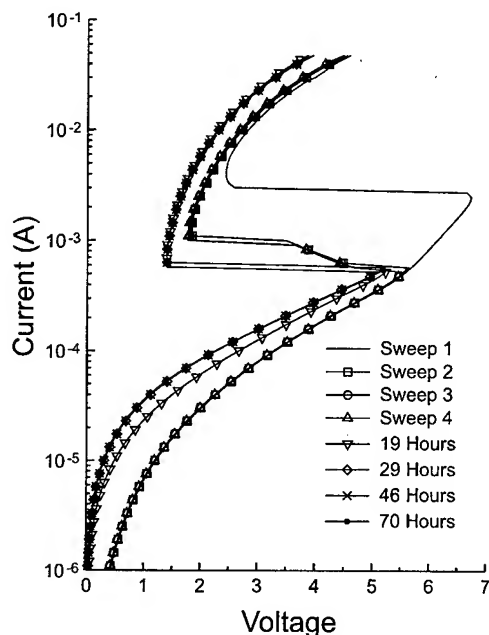


FIG. 3. Effect of stress on TSD I - V curve over time. The first four sweeps represent the "virgin" scan, taken the first time bias was ever applied to the device, and the three subsequent scans. The remaining curves were taken after the indicated number of hours of stress.

back of the chip. Discrete devices were defined by lithographically etching the top Al surface into squares of various sizes.

The device I - V curves were measured at room temperature by placing the chips on a conductive chuck of a probe station, contacting the top Al pads with a tungsten probe tip, sourcing current to the device over an exponential range, and measuring voltage. Light from a lamp directed through a microscope connected to the probe station could be injected

in the device with varying intensity in order to modify the I - V curve as shown in Fig. 2(b) and described above.

C. Results and discussion

In this section we discuss properties of the TSD I - V curves that may be related to the transport properties through the oxide layer. In particular, we explore two general characteristics of the devices: an initial "forming" behavior, and the scaling of the current with device size.

We have observed, as have others,¹⁸ that the first time a large enough bias is applied to force the device into the high current state, the I - V curve appears to change noticeably, but subsequent stressing of the device results in very little change. In Fig. 3, we show several scans taken on a single $150 \times 150 \mu\text{m}$ device: a "virgin" scan taken before any bias had been applied to the device, three additional scans taken immediately thereafter, and four subsequent scans taken after the device had been continuously ramped from negative to positive bias at 60 Hz for the indicated total number of hours. The ac stress was designed to continuously cycle the device back and forth from the high current state to the low current state in order to determine whether switching between states would cause further degradation. We would expect that degradation of the oxide would be similar for unipolar stressing as the p - n junction limits the current in the reverse bias, and therefore little voltage is dropped across the oxide. Note that the "virgin" scan exhibits a much higher peak voltage than all subsequent scans, but that no other sudden changes in the I - V curve are apparent. We suggest that the initial stress could be causing a modification of the oxide. We will return to this point shortly.

We have also observed that the I - V characteristics of different size devices have not scaled with area as would be expected if the current density were uniform. Typical I - V

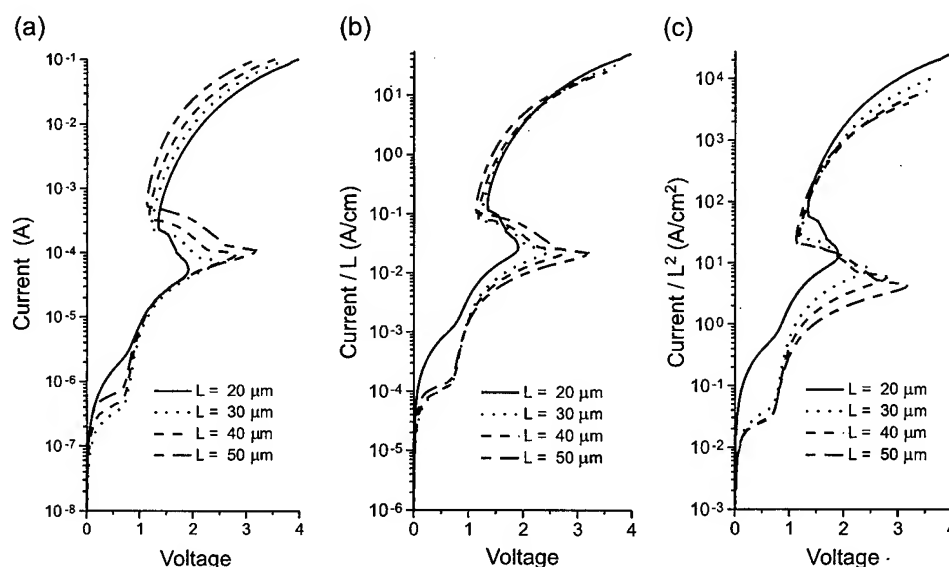


FIG. 4. Scaling behavior of the TSD device. Typical I - V curves for square devices ranging in size from $20 \times 20 \mu\text{m}^2$ to $50 \times 50 \mu\text{m}^2$ plotted as (a) current/ L^0 , (b) current/ L^1 , and (c) current/ L^2 , where L is the length of the side of the device (i.e., 20, 30, 40, and $50 \mu\text{m}$).

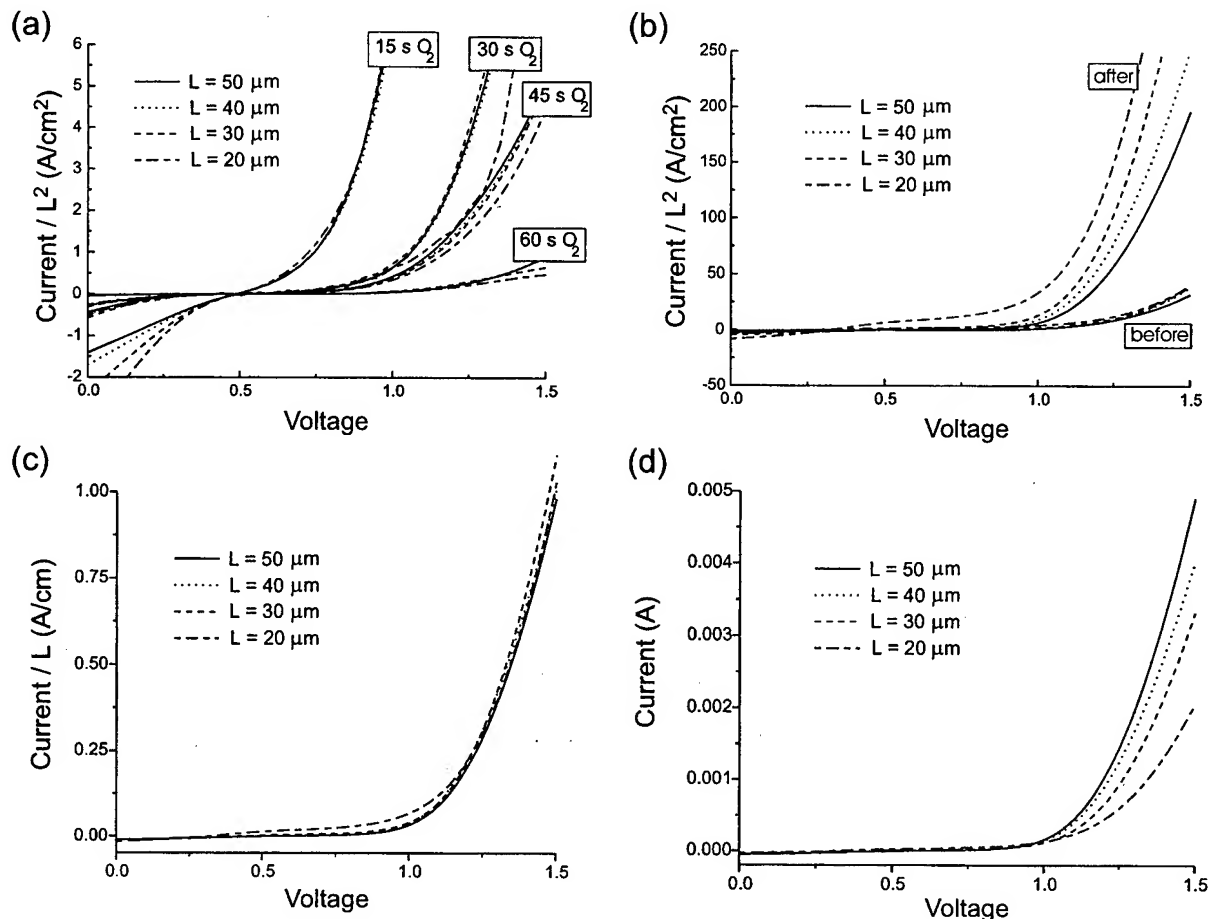


FIG. 5. Low voltage scaling behavior of the TSD with light injection. (a) "Virgin" J - V curves taken before the devices had ever been subjected to high field stress. Devices of four different sizes are shown in groups corresponding to samples with different oxide thicknesses. Each group collapses to nearly the same current density, suggesting a uniform current density. (b) J - V curves for a group of different size devices on the 30 s O_2 sample taken both before and after high field stress. The "after" curves are also plotted as (c) current/ L^1 and (d) current/ L^0 in order to illustrate the modified scaling behavior.

curves for devices of a range of sizes on the 30 s O_2 sample are shown in Fig. 4, scaled by the length (L) of the edge of the square Al contact that defines the active area of the device, raised to the powers 0, 1, and 2. If the current density through the device were uniform, the current/ L^2 curves shown in Fig. 4(c) would collapse to a single curve, but we see that different parts of the I - V curve appear to scale with different powers of L . For example, the low current region just below the peak seems to be independent of area, while the high current region above the peak seems to scale more nearly linearly with L than as any other integral power, suggesting an edge effect. It is possible that parasitic resistances could play a role in the high current region, but this would typically result in a current independent of area.

In order to determine whether the oxide plays a role in this scaling behavior, we would like to isolate its contribution to the I - V curve. As described above and illustrated in Fig. 2(c), light injection causes the peak voltage to be reduced and, if intense enough, can eliminate the peak altogether, effectively forcing the device into the high current state at any bias. In this state, the current through the device is primarily controlled by oxide tunneling as the majority of

the bias falls across the oxide. Therefore, with high enough light injection, we can effectively examine the I - V characteristic of the MOS junction even at low bias.

Using this light injection technique, we examined the MOS part of the device under low bias both before and after the device had been subjected to the stress of the high bias, high current regime that forces the "forming" process. The light level was set to completely eliminate the peak in the I - V curve. In Fig. 5(a), we plot pre-stress current density voltage (J - V) curves for four different stress current densities on each of four chips with different O_2 flow times during oxidation, and hence different oxide thicknesses. We see that the curves very nearly collapse into four groups corresponding to the four different thicknesses, suggesting nearly uniform current density through each of the oxides. While the curve corresponding to the 20 μm device on the 30 s O_2 sample does not quite coincide with the other curves for this sample, it is possible that the probe tip scraped off part of the Al pad, resulting in a smaller effective area.

In Fig. 5(b), we plot similar J - V curves of a set of different size devices on the 30 s O_2 chip under illumination both before and after the devices had been stressed. This stress

was achieved by connecting the devices to a curve tracer and, for a few seconds, switching from negative to positive bias at 60 Hz to a maximum positive bias current of 100 mA. We see, first of all, that the (average) current density through the devices increased by over an order of magnitude after stressing. We also note that the currents no longer scale with L^2 . In fact, we see by comparing Figs. 5(b), 5(c), and 5(d), that the current most closely scales linearly with L . We suggest that the oxide may be forming locally near the edges of the device where the electric field is somewhat higher. Similar behavior was observed for the 15, 45, and 60 s O_2 samples.

It is not surprising that our device geometry would produce edge effects, and it is probable that such effects could be minimized with a different device geometry, for example by extending the gate electrode across a thicker field oxide surrounding the device, or by using a guard ring. The interesting result, however, is that in the presence of the locally enhanced field around the edge of the device, the "formed" oxide is relatively stable under continued stress. As can be seen in Fig. 3, after the "forming" process, the device can be operated for long times with little modification of the peak or valley voltages at stresses far exceeding acceptable MOS values (Table I). Others have observed similar long lifetimes in these devices after an initial "forming."^{13,18} Even though this forming process greatly enhances carrier transport through the oxide locally, the oxide layer must still be functioning as an effective barrier, or the peak region of the I - V curve would disappear as described earlier.

While these results reflect a local modification of the device oxide, it is difficult to extract the oxide properties from the functions of the TSD, which is not completely well understood, in a geometry which clearly produces nonuniform transport. It would be interesting to study a simpler system with many of the same properties to see if it exhibited similar local modifications of the oxide. For these reasons, we attempted to mimic the high field local stress produced by the edge of our TSD devices with tunnel current injected from an STM tip through a thin oxide on a simple silicon substrate as described in Sec. III.

III. LOCAL PROBE STUDIES

A. Introduction

Scanning tunneling microscopy (STM) allows one to take spatially resolved measurements of the tunnel current flowing from a sharp tip to a sample surface.^{23,24} On conductive samples made of a single material, the tip may be brought to within angstroms of the surface and held at a constant voltage while the tip is raster scanned over the surface and its height modulated to maintain a constant tunnel current, effectively yielding a physical contour plot of the surface down to atomic dimensions. Similarly, on flat surfaces comprised of materials differing in conductivity, the STM may be used to generate a conductivity contour plot of the surface.^{25,26} The atomic spatial resolution of the STM makes it an attractive tool for the study of conducting and semiconducting samples. Insulating samples, on the other hand, do not gen-

erally allow a sufficient tunneling current to flow for adequate tip response resulting in frequent tip crashing, and therefore do not lend themselves as readily for STM study.

Thick layers of silicon dioxide would normally fall into the category of materials for which STM is not a useful tool. However ultrathin (<3.5 nm) layers of silicon dioxide, such as those used in the TSD, allow sufficient tunneling current to make STM a viable tool for examining local conductivity. Others have been able to study ultrathin oxides in experiments with beam assisted scanning tunneling microscopy (BASTM),²⁷ with BEEM,^{10,12,28} and with STM of thinner oxides.^{25,29} We have found that by operating at high tip voltages (1–5 V) and lower than usual tunneling currents (1–5 pA), the tip may be rastered over the silicon dioxide surface, yielding an image without crashing. The image obtained in this fashion represents some convolution of the silicon surface contour, the silicon dioxide surface contour, and any variations in the local conductivity of the silicon dioxide layer. We have attempted to characterize the local conductivity of our thin oxide layer with these STM images, comparing them with topological AFM images of our sample before and after oxidation. Furthermore, since the nature of STM is to inject large local current densities, we have also attempted to look at the effects of localized electrical stressing on the thin oxide layer.

B. Experimental details

Samples were produced using boron doped p -type (100) silicon wafers (0.00425 – $0.00575 \Omega \text{ cm}$) cleaved into square chips ~ 1 cm on a side. These chips were subject to a standard acetone, isopropyl alcohol, de-ionized water degrease in ultrasound. The chips were then etched in 10:1 HF solution to remove any oxide, and a grid of circles $2 \mu\text{m}$ in diameter, $5 \mu\text{m}$ apart, was patterned onto the chips. Chemically assisted ion beam etching (CAIBE) was employed whereby 50 sccm of chlorine gas was used to react with the bare silicon surface while a beam of xenon ions was used to sputter the chlorine-reacted silicon for 15 s. After another degrease, the bare patterned silicon surface was then examined on a Digital Instruments Nanoscope III MultiMode AFM. The chips were then degreased once again and dipped in 10:1 HF acid to remove any residual oxide. An AG Associates Mini-pulse rapid thermal Processor (RTP) was used to bring the chips up to 900°C in ultrapure argon. Device quality oxide was then grown by exposing the chips to ultrapure oxygen at 900°C for 15–30 s followed by an anneal in ultrapure argon for 30 s at 900°C . Immediately following oxidation, the chips were mounted with conductive silver paint onto magnetic pucks for use with the AFM/STM and allowed to dry for 20 min.

After drying, the thickness of the oxide layer was measured on a J. A. Woollam variable angle spectroscopic ellipsometer (VASE). Data was taken at 70° , 75° , and 80° angles of incidence across wavelengths from 205 to 365 nm. The thickness was computed using bulk SiO_2 optical constants to be 34–35 Å thick for the grown RTP oxides. AFM surface images of the samples were taken in ambient both to image

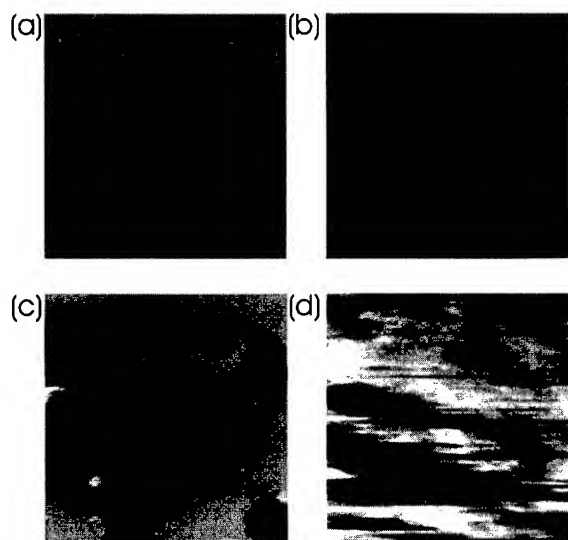


FIG. 6. Comparison of AFM and STM scans of *p*-type Si with a 34 Å RTP grown oxide. Large scale scans (10 μm , 50 nm light to dark) show the patterning of the chips in (a) an AFM scan and (c) an STM scan (4.3 V, 3.3 pA). Small scale scans (100 nm, 5 nm light to dark) taken in unpatterned areas show (b) the surface roughness in the AFM scan, and (d) the lateral variance in conductivity in the 1.0 V, 2.0 pA STM scan.

the large scale patterning (10 μm square) and the oxide surface in the unpatterned regions (100 nm square). The samples were then moved to a similar STM setup in ambient with a picoamp boost stage for use at low currents, and a variety of STM scans were taken over a period of several hours. Following study with the STM, the surface was again examined with the AFM and the oxide thickness was again measured on the ellipsometer.

We attempted to quantify the effects of atmospheric exposure on the RTP grown oxide due to native oxide growth and accumulation of atmospheric moisture or other contaminants. Unpatterned samples were degreased, etched, and oxide layers were grown in the RTP in a similar fashion to the patterned samples. Timed thickness measurements taken over periods of 10–15 h in ambient exhibited linearized growth rates on the order of 0.2 Å/h. This correlates well with the ellipsometric observed change in thickness of several angstroms over the time of the STM scanning.

C. Results and discussion

The AFM was used to examine the topological contour of chips both before and after oxidation. Before oxidation, the bare silicon surface was examined both to verify that the patterning was successful and to quantify the silicon surface roughness. Section linescans of the plots showed the silicon surface to have local roughness on the order of 1 nm. Post-oxidation Si–SiO₂ interface roughness has been found to be less than 1 nm by others.³⁰ After RTP oxidation, the AFM was again used to image the patterning and surface roughness. Figure 6 shows typical AFM scans of both the large scale patterning and small scale roughness of the RTP thin oxide layer. Figure 6(a) shows the large scale patterning image, found to have patterned hole depths of 18 nm. Figure

6(b) shows a 100 nm scan of the RTP oxide surface roughness. Line scans of the 100 nm image show small scale roughness to be less than 1 nm in the unpatterned areas of the chip.

The STM was then used to look at the electrical contour of the RTP thin oxide layer. An initial attempt was made to image the surface in a constant-height mode, looking only at local changes in tunneling currents without allowing the tip to respond to the surface. However, the very dramatic changes in current across our oxide resulted in extreme noise, frequent tip crashes, and an irreproducible image. In constant-current mode, we are assured that if the sample is uniformly conductive, the surface image will reproduce that of the AFM. By comparing our image with the AFM, we are able to distinguish differences in conductivity in the oxide layer from topological features. The STM images were taken at relatively high voltages (1–5 V) and low currents (1–5 pA) to keep the tip sufficiently far away from the surface to avoid crashing and also to drive sufficient tunnel current through the oxide for adequate feedback. This high voltage low current imaging was done at the expense of lateral resolution, but was necessary to generate repeatable scans. Similar setpoint voltages with nanoamp currents have been used by others in STM experiments with thinner oxides,^{25,29} as well as BASTM,²⁷ and BEEM^{10,12,28} experiments.

Initially the chips were scanned over 10 μm to image the patterned circles to verify that our scan parameters were producing a reliable image. Figure 6(c) shows a typical large scale STM image (4.3 V, 3.3 pA) of the patterned surface on the same scale as the corresponding AFM image in Fig. 6(a). The images of the patterned circles show depths of 18.5 nm in excellent agreement with the AFM scan, well within expected variation in etch depth across the sample. Having established that the STM is producing a reliable image of the oxidized chip, the scan size was reduced to look at local variations in the oxide. Figure 6(d) shows a 100 nm STM image (1.0 V, 2.0 pA) of the oxide layer on the same scale as the corresponding AFM image in Fig. 6(b). Line scans show apparent height variations on the order of 5 nm across areas 20–50 nm in size. The dramatic difference in the images taken with AFM and STM, coupled with the AFM verified flat topology, suggest that there are significant lateral variations in conductivity across the oxide. In order to verify that these images represented true surface features, we attempted to reproduce them with a variety of scan parameters. Surface features were relatively reproducible under changes in scan rate, direction, and area. Figure 7 shows a representative series of STM images (1.5 V, 3.0 pA) ranging in size from 1 μm to 100 nm on the same unpatterned section of RTP oxide. Apparent height variations in all four images are on the order of 5 nm with feature size on the order of 20–50 nm laterally in all images. Lateral variations in conductivity similar in size and magnitude were seen across a number of different scans and chips.

Note that in the interpretation of these variations as conductive rather than topological features, we require a tip-sample separation of at least 5 nm. This is considerably

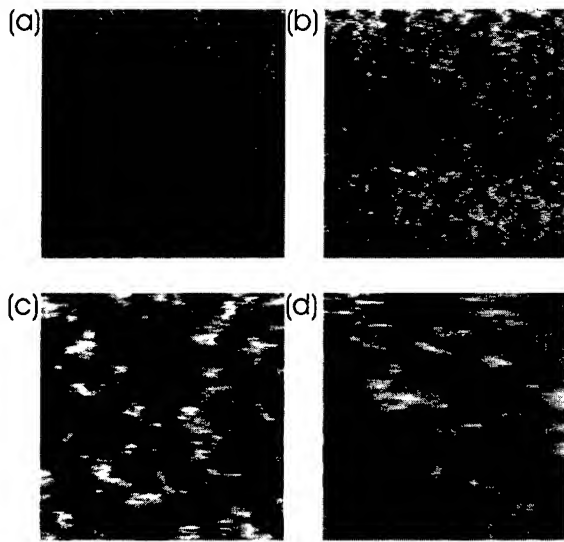


FIG. 7. Zoom in sequence of STM scans (1.5 V, 3.0 pA) over an unpatterned area of 34 Å RTP oxide on *p*-type Si. Lateral variations in conductivity can be seen over the sequence of scans: (a) 1 μ m, (b) 500 nm, (c) 250 nm, and (d) 100 nm scans. Light to dark variation in the scans is 5 nm.

larger than the distances typically assumed in vacuum STM setups, and would suggest that the tip current in our experimental setup is considerably less sensitive to tip-sample distance. This is a reasonable assumption for ambient STM; accumulation of moisture and atmospheric contaminants on the tip and sample due to the high fields present can serve to reduce the sensitivity of current to tip-sample distance.³¹ Similar larger than expected tip-sample distances have been previously reported in other STM experiments in ambient.²⁵

The STM was also used to look at the effect of large currents on localized areas of the chip. Two different methods of stressing were employed. In one stress mode, the tip was held at one spatially localized area, and an *I-V* curve was taken, causing the voltage to be ramped repeatedly from positive to negative bias. After scanning a 10 μ m area, an *I-V* curve was taken over a bias range of ± 5 V or more. Upon returning to the imaging mode (1.5 V, 1–2 pA), a large dark area hundreds of nanometers in size and more than 30 nm deep was present. A similar effect was noted in another stress mode, whereby scans were kept at imaging voltages and currents but concentrated in lateral dimension. Figure 8

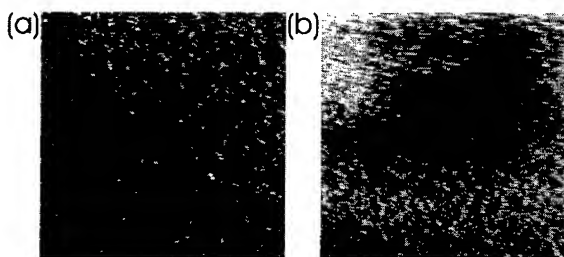


FIG. 8. (a) Before and (b) after 1.5/1.0 V, 3.0 pA STM scans of a 1 μ m, 20 nm light to dark, area subjected to a zoom in sequence of scans. The 10 nm deep "hole" visible in (b) after zoom in represents an area of reduced conductivity.

illustrates the appearance of a modified surface feature after a series of scans similar to those in Fig. 7. Figure 8(a) shows an initial scan of a 1 μ m area, and Fig. 8(b) a rescan of the same 1 μ m area after 500, 250, and 100 nm scans had been performed within it. The scan of the modified surface contains a large dark area 10 nm deep where the smaller, more concentrated, scans had occurred, similar to that which was present after the *I-V* scans. These features are likely due to the creation and charging of trap states which reduce the conductivity of the sample, as seen by others.^{12,28}

IV. DISCUSSION

It is difficult to conclusively compare the results of these two sets of experiments. It is unclear whether the small scale lateral variations seen in the STM images on the unstressed oxides could play a role in the "forming" behavior of the devices or the ability of the formed oxide to continue to act as an effective barrier. It is possible that stress modification of the different conductivity regions could be related to the observed "forming" behavior, but unfortunately, due to the large drift in the lateral tip position, we have as yet been unable to selectively stress the different conductivity regions with the STM in order to test this hypothesis.

While we observed an increase in the oxide conductivity in the stressed regions of the TSD device, we observed a decrease in the oxide conductivity with stress in the STM experiments. Others¹¹ have observed an initial decrease in conductivity followed by an increase in conductivity with continued charge injection in BEEM stressing experiments. Assuming a similar behavior in our system, is possible that we were unable to inject a high enough charge density to stress the oxide to the point of increasing conductivity due to tip drift or current spreading.

V. CONCLUSIONS

Observations of the scaling behavior of the TSD device indicate that under initial heavy electrical stress, the silicon dioxide layer appears to become more conductive in the vicinity of the edges of the gate electrode where the electric field is expected to be highest. It does not break down in the conventional sense, however, as it still acts as enough of a tunnel barrier that the device supports multistable current states. AFM and STM measurements on similarly grown oxides indicate that, while the silicon and oxide surfaces are quite flat topologically, there are significant lateral electrical variations, even before significant stress has been applied. Stressing the oxide by injecting large currents with large voltages, or by concentrating a raster scan over a small area, results in an area of lower conductivity in the STM image, likely the result of negative charge trapping in the oxide.

ACKNOWLEDGMENTS

This research was supported in part by the Defense Advanced Research Projects Agency and monitored by the Office of Naval Research under Contract Number N0014-93-1-0710.

- ¹D. A. Buchanan and S.-H. Lo, *Proc. Electrochem. Soc.* **96-1**, 3 (1996).
- ²K. F. Schuegraf and C. Hu, *J. Appl. Phys.* **76**, 3695 (1994).
- ³R. Rofan and C. Hu, *IEEE Electron Device Lett.* **12**, 632 (1991).
- ⁴M. Hirose, J. L. Alay, T. Yoshida, and S. Miyazaki, *Proc. Electrochem. Soc.* **96-1**, 485 (1996).
- ⁵T. Yoshida, D. Imafuku, J. L. Alay, and S. Miyazaki, *Jpn. J. Appl. Phys.* **1** **34**, L903 (1995).
- ⁶M. O. Andersson, K. R. Farmer, and O. Engstrom, *J. Appl. Phys.* **71**, 1846 (1992).
- ⁷K. R. Farmer, M. O. Andersson, and O. Engstrom, *Appl. Phys. Lett.* **60**, 730 (1992).
- ⁸Z. H. Lu, M. J. Graham, D. T. Jiang, and K. H. Tan, *Appl. Phys. Lett.* **63**, 2941 (1993).
- ⁹A. Pasquerello, M. S. Hybertsen, and R. Car, *Phys. Rev. B* **53**, 10942 (1996).
- ¹⁰R. Ludeke, H. J. Wen, and E. Cartier, *J. Vac. Sci. Technol. B* **14**, 2855 (1996).
- ¹¹R. Ludeke, A. Bauer, and E. Cartier, *Appl. Phys. Lett.* **66**, 730 (1995).
- ¹²B. Kaczer and J. P. Pelz, *J. Vac. Sci. Technol. B* **14**, 2864 (1996).
- ¹³T. Yamamoto, K. Kawamura, and H. Shimizu, *Solid-State Electron.* **19**, 701 (1976).
- ¹⁴N. T. Ali and R. J. Green, *IEEE Trans. Electron Devices* **42**, 1978 (1995).
- ¹⁵J. G. Simmons and A. A. El-Badry, *Radio Electron. Eng.* **48**, 215 (1978).
- ¹⁶H. J. Levy and T. C. McGill, Patent No. 5,535,156 (1996).
- ¹⁷P. O. Pettersson, A. Zur, E. S. Daniel, H. J. Levy, O. J. Marsh, and T. C. McGill, *IEEE Trans. Electron. Dev.* (submitted).
- ¹⁸H. Kroger and H. A. R. Wegener, *Solid-State Electron.* **21**, 643 (1978).
- ¹⁹H. Kroger and H. A. R. Wegener, *Solid-State Electron.* **21**, 655 (1978).
- ²⁰A. M. Hodge, C. Pickering, A. J. Pidduck, and R. W. Hardeman, *Mater. Res. Soc. Symp. Proc.* **52**, 313 (1986).
- ²¹S. E. Lassig, T. J. Debolske, and J. L. Crowley, *Mater. Res. Soc. Symp. Proc.* **92**, 103 (1987).
- ²²S. Duzelier, G. Sarabayrouse, J. L. Prom, and G. Hollinger, *Electron. Lett.* **27**, 1399 (1991).
- ²³J. Tersoff and D. R. Hamann, *Phys. Rev. Lett.* **50**, 1998 (1983).
- ²⁴J. Tersoff and D. R. Hamann, *Phys. Rev. B* **31**, 805 (1985).
- ²⁵R. Chapman, M. Kellam, S. Goodwin-Johansson, J. Russ, and G. E. McGuire, *J. Vac. Sci. Technol. B* **10**, 502 (1992).
- ²⁶R. M. Feenstra, *Semicond. Sci. Technol.* **9**, 2157 (1994).
- ²⁷S. Heike, Y. Wada, M. Lutwyche, K. Murayama, and H. Kuroda, *Jpn. J. Appl. Phys.* **1** **34**, 1376 (1995).
- ²⁸B. Kaczer, Z. Meng, and J. P. Pelz, *Phys. Rev. Lett.* **77**, 91 (1996).
- ²⁹M. Tabe and M. Tanimoto, *Appl. Phys. Lett.* **58**, 2105 (1992).
- ³⁰M. Niwa and H. Iwasaki, *Jpn. J. Appl. Phys.* **1** **28**, 2320 (1989).
- ³¹J. Jahanmir, P. E. West, A. Young, and T. N. Rhodin, *J. Vac. Sci. Technol. A* **7**, 2741 (1989).

Plasma-assisted formation of low defect density SiC-SiO₂ interfaces

A. Götz

Institut für Halbleitertechnik, RWTH-Aachen, Aachen D-52074, Germany

G. Lucovsky,^{a)} K. Koh, D. Wolfe, and H. Niimi

Departments of Physics, Materials Science and Engineering and Electrical and Computer Engineering, and Chemistry, North Carolina State University, Raleigh, North Carolina 27695-8202

H. Kurz

Institut für Halbleitertechnik, RWTH-Aachen, Aachen D-52074, Germany

(Received 24 February 1997; accepted 22 April 1997)

The initial stages of SiC-SiO₂ interface formation by low temperature (300 °C) remote plasma assisted oxidation (RPAO) on flat and vicinal 6H SiC(0001) wafers with Si faces have been studied by on-line Auger electron spectroscopy (AES). Changes in AES spectral features associated with Si-C and Si-O bonds are readily evident as oxidation progresses; however, there are no detectable AES features that can be attributed to C-O bonds. Initial oxidation rates as determined from AES data are greater for vicinal wafers than for flat wafers paralleling results for RPAO oxidation of Si. Devices fabricated on vicinal SiC wafers require an 1150 °C anneal in an H₂ containing ambient to reduce defect densities from the 10¹³ to 10¹¹ cm⁻² range, consistent with termination of C atom step edge dangling bonds by H atoms. Devices prepared by thermal oxidation also require a 1150 °C anneal in H₂ even though silicon oxycarbide regions with C-O bonds are formed in a transition region at the SiC-SiO₂ interfaces. © 1997 American Vacuum Society. [S0734-211X(97)07304-6]

I. INTRODUCTION

There is considerable interest in SiC for high temperature device applications, including power transistors. To fabricate these devices, it will be necessary to gain an increased understanding of the chemical bonding at the SiC-SiO₂ interface, and in particular to develop processing that minimize interfacial defects and transition regions, and maximize device reliability as manifested in reductions in the rate of interfacial and bulk defect generation during operation. It has been established that the oxide that forms on SiC during high temperature oxidation in O₂ ambients is SiO₂, and that this oxidation process can yield SiC-SiO₂ interfaces and gate dielectrics with sufficiently low defect densities for field effect transistor (FET) operation.¹ However, channel mobilities at room temperature (in depletion mode FETs) are significantly lower than expected from bulk SiC properties and much research has been focused on explaining these differences and improving channel mobilities.² Since SiO₂ is the dominant solid state oxidation product, oxidation products involving carbon atoms must include gaseous molecules such as CO; as oxide growth progresses these must be transported from the buried growth interface through the oxide and out of the film. However, this does not preclude the formation of additional interfacial carbon atom bonding arrangements such as Si-C and C-O as have been reported in silicon oxycarbide transition regions.³ The molar volume mismatch between SiC and SiO₂ is greater than the corresponding mismatch between Si and SiO₂, so that intrinsic levels of strain at SiC-SiO₂ interfaces are anticipated to be at least as high as the values of approximately 5 × 10⁹ dynes/cm² reported in Si-SiO₂.⁴ This article deals

with an alternative low temperature method for semiconductor-dielectric interface formation that has been successfully applied to crystalline Si device technology.^{5,6} As applied to Si, this approach provides independent control of Si-SiO₂ interface formation and dielectric film growth. The Si-SiO₂ interface and a superficially thin oxide layer (~0.5 nm) are formed by a low temperature (300 °C) remote plasma-assisted oxidation (RPAO) process, and the remainder of the oxide layer is deposited by a low temperature (300 °C) remote plasma enhanced chemical vapor deposition (RPECVD) process. Device quality interfaces are obtained by subjecting the plasma processed Si-SiO₂ heterostructure to a post-deposition rapid thermal anneal (RTA), e.g., for 30 s at 900 °C. The 900 °C anneal promotes both chemical and structural interface relaxations, including a minimization of interfacial sub-oxide bonding arrangements in transition regions between the Si crystal and the stoichiometric bulk oxide.⁷ A similar plasma-assisted process has recently been applied to SiC, and has yielded SiC-SiO₂ interfaces with defect state densities comparable to what has been achieved on SiC by conventional high temperature thermal oxidation.⁸ The potential advantages of this two step approach to SiC-SiO₂ interface formation are: (i) it reduces the requirement for transporting gaseous reaction products such as CO molecules away from the interface as the oxide grows; and (ii) it can also reduce strain induced interface defect generation, since only a small fraction of the total oxide layer thickness is generated by consumption of the SiC substrate. The focal point of this article is a study of SiC-SiO₂ interface formation by RPAO using on-line Auger electron spectroscopy (AES) to monitor interface bonding chemistry and the initial oxide growth rate.

Before discussing the SiC AES measurements, it is useful

^{a)}Electronic mail: gerry_lucovsky@ncsu.edu

to summarize the results of device studies on vicinal Si and SiC wafers using the plasma-assisted processing steps identified above. Studies by Bjorkman *et al.*⁹ demonstrated that densities of mid-gap defect states (D_{it}) as low as $1 \times 10^{10} \text{ cm}^{-2} \text{ eV}^{-1}$ could be obtained on Si(111) surfaces by using conventional thermal oxidation, metal electrodes and 400 °C post metal anneals in an H₂ containing ambient. Previous studies had always shown higher values for Si(111) as compared to the Si(100) surfaces used in the integrated circuit wafers (see Reference 9). Studies performed by Yasuda *et al.* on *p*-type Si(111) vicinal up to 5° in the 11 $\bar{2}$ direction showed that the density of mid-gap states (D_{it}) increased by relatively small amounts from $\sim 1 \times 10^{10} \text{ cm}^{-2} \text{ eV}^{-1}$ to about $6 \times 10^{10} \text{ cm}^{-2} \text{ eV}^{-1}$ as the vicinal angle increased to about 5°. The studies of Yasuda *et al.* were on Si–SiO₂ structures prepared by the same two-step RPAO-RPECVD 300 °C process that has been described above, and is discussed in detail in Ref. 5. Al electrodes were employed and a 400 °C postmetallization anneal (PMA) in a H₂ containing ambient was used. The magnitudes of the incremental increases in D_{it} with increasing vicinal angle indicated that only a relatively small fraction of the dangling bond states at the step edges contributed to increases in D_{it} with increasing step edge density due to increasing vicinal angle.

At the present state of the art, the highest quality SiC crystals for devices are epitaxial layers formed on Si atom terminated 6H SiC vicinal approximately 3.5° in the 11 $\bar{2}$ direction. Using RPAO-RPECVD, the results of device studies using vicinal 6H *p*-type SiC as reported by Gölz *et al.*⁷ are markedly different than the results reported for vicinal Si. Metal-oxide-semiconductor (MOS) capacitors were prepared the essentially the same two-step RPAO-RPECVD process as used by Yasuda *et al.*⁵ for Si. However, following a conventional PMA in an H₂ containing ambient, D_{it} values were in excess of $10^{13} \text{ cm}^{-2} \text{ eV}^{-1}$, even when an 1150 °C furnace anneal in Ar was interposed between the oxide film deposition and the metallization steps. However if the high temperature anneal was performed in an H₂ containing ambient then D_{it} values after the conventional 400 °C PMA were reduced significantly into the $10^{11} \text{ cm}^{-2} \text{ eV}^{-1}$ range.⁸ This large difference in defect densities is approximately equal to the number of carbon atom dangling bonds at the step edges suggesting that these dangling bonds were not H-terminated after a conventional 400 °C PMA, but instead required the higher temperature exposure to H for termination. It is important to note that interfaces prepared on the same type of vicinal SiC surfaces by conventional high temperature thermal oxidation processes also required a high temperature anneal in an H₂ containing ambient to yield values of D_{it} in the $10^{11} \text{ cm}^{-2} \text{ eV}^{-1}$ range.¹⁰ Step edge carbon atom dangling bonds on vicinal Si wafers can either be terminated by oxygen atoms that connect into the SiO₂ dielectric, or by hydrogen atoms. The annealing results in Refs. 8 and 10 are consistent with hydrogen atom termination of step edge carbon atom dangling bonds (see Section II for atomic structure at the step edges); i.e., dangling bond elimination

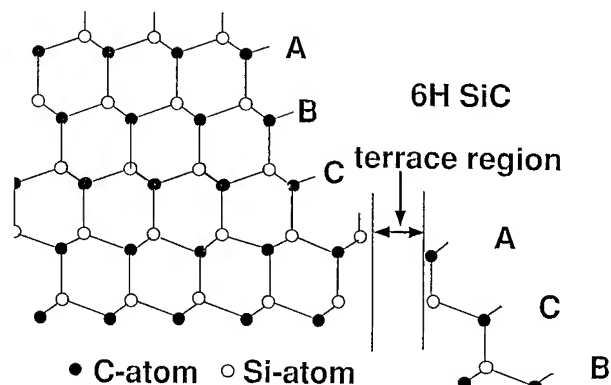


FIG. 1. Schematic representation of surface and step bonding for 6H SiC vicinal wafers. The C atom dangling bonds for the first 6-atom, 3-layer sequence are one per step edge C atom; for the second 6-atom, 3-layer sequence, there are two dangling bonds per step edge C atom. The 6-atom steps are separated by flat terrace regions with Si-atom dangling bonds.

would be accomplished by atomic hydrogen generated during the 1150 °C anneal.

To study interfacial bonding arrangements between SiC and SiO₂, on-line AES has been used to monitor the initial stages of SiC–SiO₂ interface formation. This same approach has been used to study the initial stages of interface formation on Si by both RPAO and rapid thermal oxidation (RTO), and using both O₂ and N₂O oxygen atoms sources gases.⁷ This article extends this approach to SiC and focuses on the initial stages of the RPAO process as applied to both flat or vicinal SiC surfaces using O₂ source gases. Preliminary results are also presented for RPAO using N₂O as the oxygen atom source gas. Section II discusses briefly the local bonding arrangements at the step edges of the vicinal 6H SiC wafers. Section III presents results of AES studies of SiC RPAO, and includes comparisons between the RPAO of SiC and Si. Detailed results are presented for RPAO using O₂, and preliminary results for RPAO using N₂O. Section IV summarizes the major results of this study as they apply to (i) the results of the on-line AES studies, and (ii) the passivation of step edge dangling bonds as revealed by changes reductions in D_{it} levels as determined from electrical measurements.

II. LOCAL BONDING AT STEP EDGES

The dangling bond geometry for Si(111) wafers off cut in the 11 $\bar{2}$ direction has been discussed in Ref. 4 and references therein. The steps are effectively two atom layers high, and the step edge atoms have a single dangling bond. The steps on 6H SiC with Si terminated surfaces are more complex due to the larger unit cell. The repeat pattern of 6H SiC extends over 12 atomic planes of alternating Si and C atoms as shown in Fig. 1. For off cut angles in the 11 $\bar{2}$ 0 direction, the vicinal Si face surfaces can either include two different 6 atom steps that are separated by Si atom terminated terrace regions, or alternatively consist of six 2 atom layer steps/6H repeat pattern. Figure 1 indicates the bonding arrangements for the two different 6 atoms steps. The first set of steps are in an ABC sequence with one carbon atom dangling bond

per double layer Si-C component, and the second set of steps are an ACB sequence with two carbon atom dangling bonds per double layer Si-C component. The average number of dangling bonds per step in 6H SiC(0001) is 1.5 dangling bonds per step edge atom, whereas in Si(111) off cut in the 112 direction it is only 1 dangling bond per step atom. The spacing between single steps on Si(111) off cut in the 112 direction is ~ 5 nm for a 3.5° off cut angle, so that the ratio of step edge to terrace dangling bonds is approximately 0.06. For a 3.5° off cut angle on SiC, the dangling bond ratio is higher by a factor of 1.5 due to the different step edge terminations of the ABC and ACB sequences. If the 6H SiC surface has two atom steps, as is likely for the wafers used in this study, the dangling bond considerations are the same. Starting with the same ABC sequence as in Fig. 1, the first three 2 atom steps are terminated by C atoms with one dangling, and the second three 2 atom steps by C atoms with two dangling bonds. This means that the number of dangling bonds per step on the average is still 1.5 times greater than for off cut Si wafers.

III. EXPERIMENTAL RESULTS

The samples used in this study were *p*-type 6H SiC purchased from Cree Corporation with Si faces. Similar to polar faces on III-V compounds, the polar faces on SiC are either Si or C faces. Note further that this particular designation is used independent of any surface contamination. One of the wafers was oriented on a principal axis, the (0001) direction, and the other was a vicinal wafer off cut at approximately 3.5° in the 1120 direction. The surface of the on-axis sample was prepared from and bulk ingot, and the surface of the off-axis sample used in these studies was from an epitaxially grown film. The surfaces of the wafers were cleaned following a procedure suggested by Cree Research, Inc.¹¹ The *ex situ* cleaning consisted of (i) sequential rinses in tri-chloroethane (TCE), acetone and methanol, followed by (ii) a conventional two bath RCA clean. The sacrificial oxide formed in the second step of the RCA clean was removed by rinse in dilute HF. Samples were then loaded into a multichamber system, which provides a separate ultrahigh vacuum (UHV) compatible chambers for RPAO and AES. The experimental procedure was to alternate AES measurements with RPAO processing. A similar approach has been applied to RPAO of Si, and through analysis of the AES data, two kinds of information were obtained: (i) the chemical bonding at the Si-SiO₂ interface, and (ii) the oxide thickness as a function of oxidation time.¹²

Figures 2(a) and 2(b), respectively give differential AES spectra for flat and vicinal SiC obtained by RPAO using an O₂ source gas. The experimental processing conditions are:

- (i) a substrate temperature of 300 °C,
- (ii) a process pressure of 300 mTorr,
- (iii) a plasma power to the He/O₂ mixture of 30 W, and
- (iv) flow rates 200 standard cubic centimeters per minute (sccm) of He, and 20 sccm and O₂.

The as-loaded samples show surface contamination by oxy-

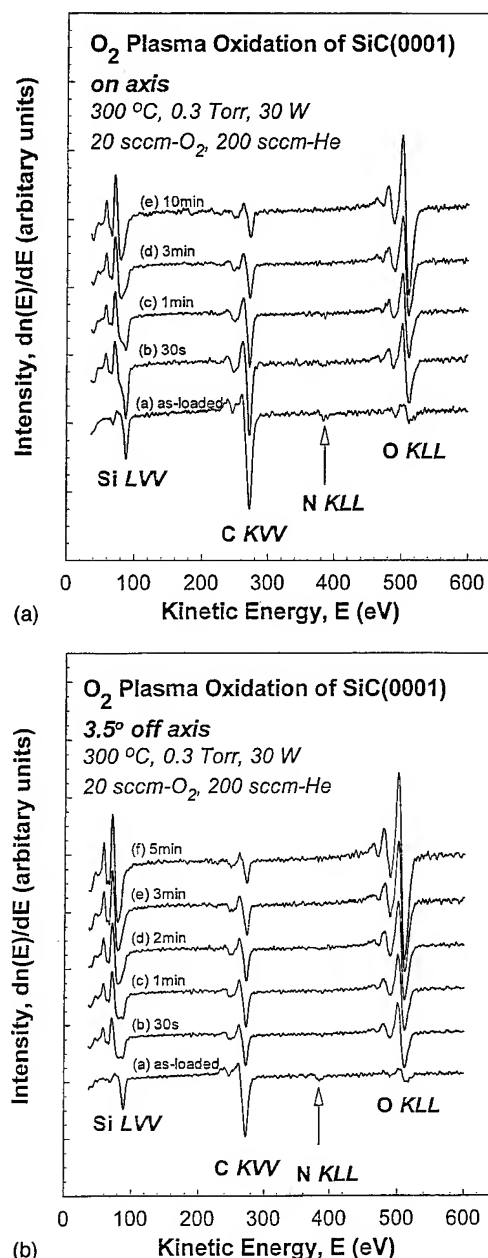


FIG. 2. Derivative mode AES spectra for (a) flat and (b) vicinal SiC obtained by RPAO at 300 °C using O₂ as the oxygen atom source gas.

gen and nitrogen that was not removed by the *ex situ* cleaning process. Since the levels of surface contamination were sub-monolayer, no *in situ* cleaning was attempted. This was in part due to the limited temperature capabilities of the plasma-processing chamber; the substrate heater in this chamber cannot heat a wafer to a temperature greater than about 500 °C. Never-the-less, the nitrogen contamination problem must be overcome before detailed studies of interface bonding using an N₂O source gas can be definitive with respect to any preferential nitrogen atom incorporation at the SiC-SiO₂ interface. It should be further noted that the SiC samples used in this study were *n*-type with a N atom doping level of $\sim 1.4 \times 10^{18} \text{ cm}^{-3}$, so that the N contamination

may have its origin in the doping of the SiC wafers; however, it cannot be attributed simply to the doping because the area density of dopant atoms in a 2 nm thickness (~ 2 electron escape depths) is $\sim 3 \times 10^{11} \text{ cm}^{-2}$, which is considerably less than the minimum AES detection limit of $\sim 10^{13} \text{ cm}^{-2}$. In spite of the sub-monolayer surface N concentration, information relative to oxidation rates could be obtained using the N₂O source gas. In the case of Si, as-loaded samples generally showed small levels of oxygen contamination, but never showed any nitrogen signal, so that more detailed interface bonding studies could be performed with both O₂ and N₂O source gases.¹² Prior to oxidation, the flat and vicinal SiC wafers displayed essentially the same ($\pm 3\%$) Si/C surface ratios so that changes in these ratios could be correlated with the effects of the oxidation processes. As the oxidation time for the SiC wafers is increased, there are four changes evident in the SiC spectra in Figs. 2(a) and 2(b):

- (i) the line shapes and multiplicity of features within the Si_LVV group changed,
- (ii) the C_KVV signal strength decreased,
- (iii) the N_KLL signal strength decreased, and
- (iv) the O_KLL signal strength increased.

These spectral changes are consistent with the growth of an SiO₂ film on the SiC substrate.

Figures 3(a) and 3(b), and 4(a) and 4(b) display, respectively, changes in (a) the Si_LVV and (b) the C_KVV (AES) features as a function of the oxidation time for flat (Fig. 3) and vicinal (Fig. 4) surfaces. Consider first the Si_LVV features in Figs. 3(a) and 4(a). The feature at $\sim 88 \text{ eV}$ is associated with Si-C bonds, in particular with a Si atom that is bonded to four C atoms. As the oxidation proceeds this feature decreases in strength as the Si-O feature at $\sim 76 \text{ eV}$ increases. These changes in relative intensity are due to increase of the oxide thickness with time and are consistent with the relative values of the electron escape depth, $\sim 0.6 \text{ nm}$, and the oxide thickness. There is also a shift of the Si-C feature to lower energy as the oxidation proceeds; this is more easily seen in the spectra for the off-axis sample. Figures 5(a) and 5(b) display the integrated Si_LVV spectra. It is evident from Figs. 3(a) and 5(a) that as the intensity of the Si-C feature decreases, the position of this spectral feature moves to lower energy. In particular in Fig. 3(a), the Si-C feature is evident as a distinct spectral peak at $\sim 88 \text{ eV}$ prior to oxidation, after the oxidation has progressed for 3 min it has shifted to lower energy and appears as shoulder at $\sim 84 \text{ eV}$ on the 76 eV peak. In contrast, the C-Si feature in the C_KVV spectrum in Fig. 4(a) simply decreases in strength as the oxidation process proceeds. The absence of any significant spectral change in this region of the AES spectrum is indicative of the fact that C-O bonds are not formed during the RPAO process. The development of interfacial C-O bonds would produce a satellite peak at lower energy, which would increase in relative strength to the C-Si feature as the oxide growth proceeds and the AES signal becomes more sensitive to the SiC interface than the bulk.

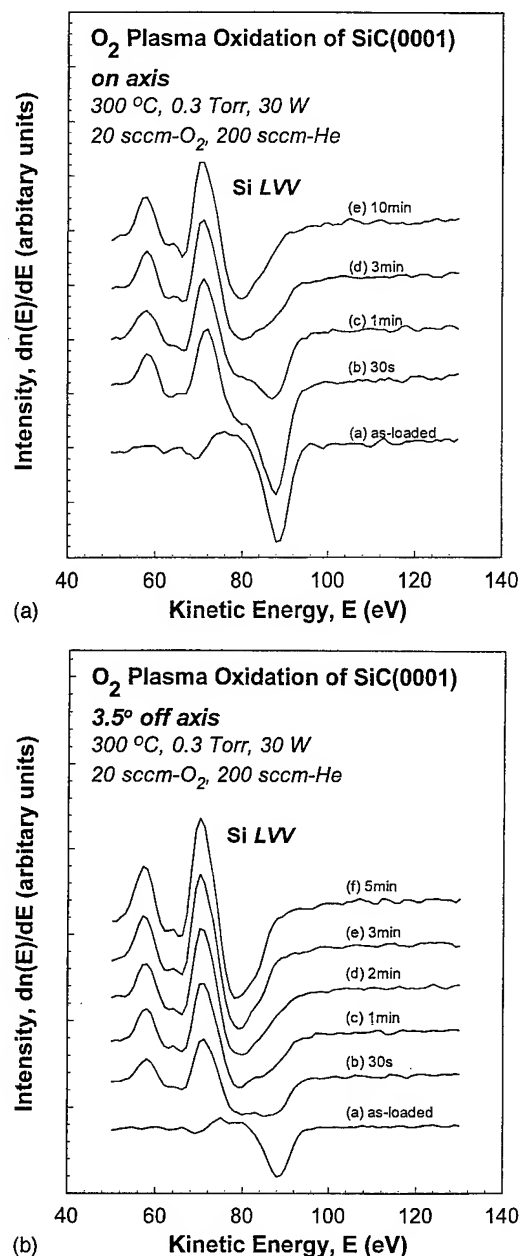


FIG. 3. Changes in the Si_LVV AES features as a function of the oxidation time for (a) flat and (b) vicinal SiC surfaces for O₂ RPAO.

The oxide thickness can be obtained to $\pm 5\%$ from the relative intensity changes in the C_KVV feature using the characteristic escape depth of 0.96 nm for 275 eV electrons. A similar approach has been applied for the determination of the oxide thickness for the RPAO process on Si, and has been validated by a direct measurement of oxide thickness by cross-sectional high resolution transmission electron microscopy.

Figures 6(a) and 6(b) shows log-log plots, respectively of the oxide thickness versus the oxidation time for the RPAO of (a) Si and (b) SiC for the O₂ source gas. In each instance the data can be fit by a power law dependence of the form, $t_{\text{ox}} = At^b$, where t_{ox} is the oxide thickness in nm, t is the oxidation time in minutes, and A and b are fit

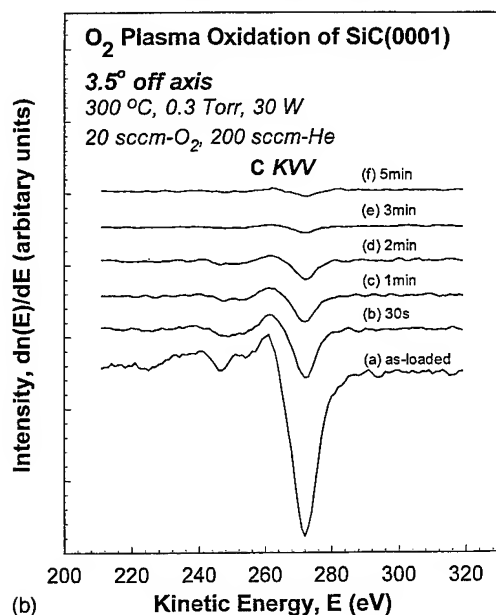
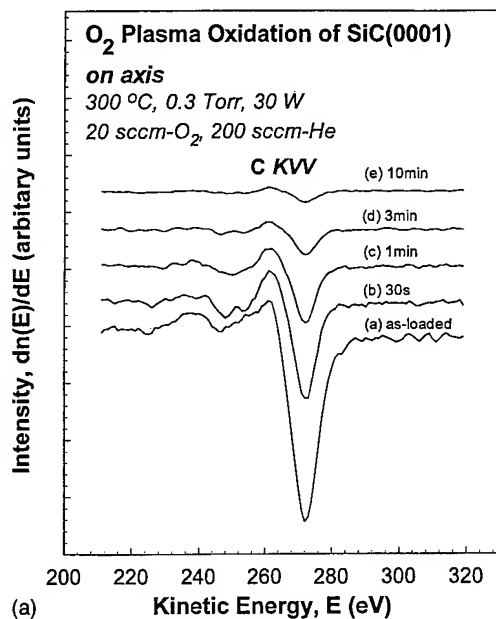


FIG. 4. Changes in the C_{KVV} AES features as a function of the oxidation time for (a) flat and (b) vicinal SiC surfaces for O₂ RPAO.

parameters.^{12,13} The initial oxidation rates are about one and one-half times faster for the vicinal than for the flat wafers, and the initial oxidation rates are faster for Si than for SiC. However, the power law exponents for the SiC oxidations for both flat and vicinal surfaces are greater than the corresponding exponential factor for Si; however, all of these power law exponents are less than 1. The relative differences in the power law factors means that at long times of about 10 min or more the thicknesses of the oxides formed on both Si and SiC are comparable for both the flat and vicinal wafers. Figure 7 illustrates this by showing a comparison of oxide thickness versus oxidation time for flat Si and SiC; a similar situation also prevails for the oxidation of the off-axis wafers.

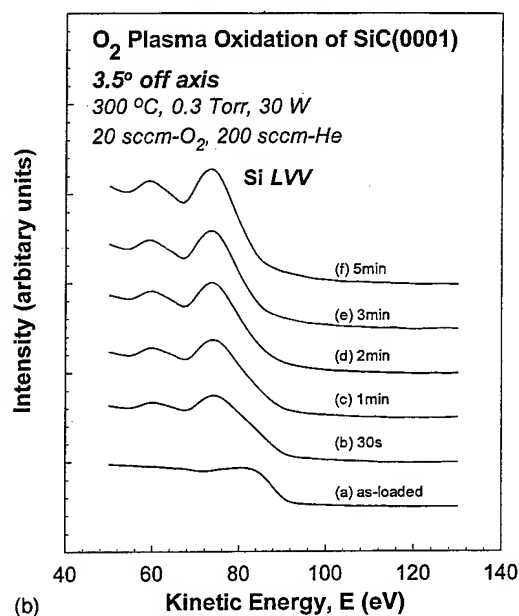
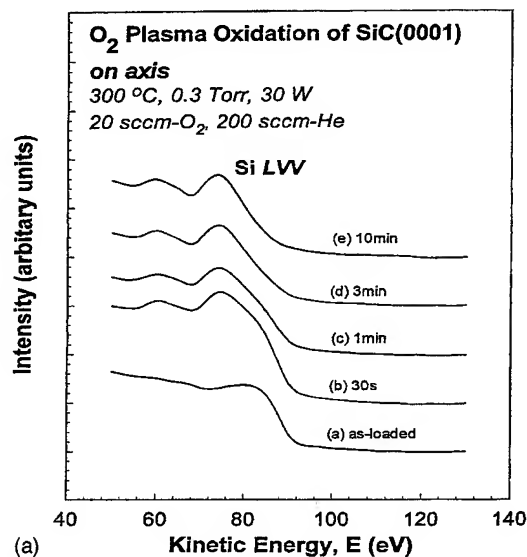


FIG. 5. Integral, $n(E)$, Si_{LVV} AES spectra as a function of the O₂ RPAO oxidation time for (a) flat and (b) vicinal SiC surfaces.

Finally for the RPAO of Si using the same relative flow rates of N₂O and O₂, the oxide growth rate was slower by a factor of 2 using the N₂O source gas.¹² Comparisons of RPAO rates have also been made for SiC for plasma excited O₂ and N₂O source gases. For the same relative flow rates of these two source gases, the situation is reversed with respect to Si, i.e., the oxidation rate of SiC is greater using the N₂O source gas. This increased growth rate using plasma excited species from the N₂O discharge is illustrated in Fig. 8 for the RPAO of flat SiC wafers.

IV. DISCUSSION AND SUMMARY

A. AES results

The analysis of the AES spectra for the flat and vicinal SiC surface RPAO processes using the O₂ source gas indicates:

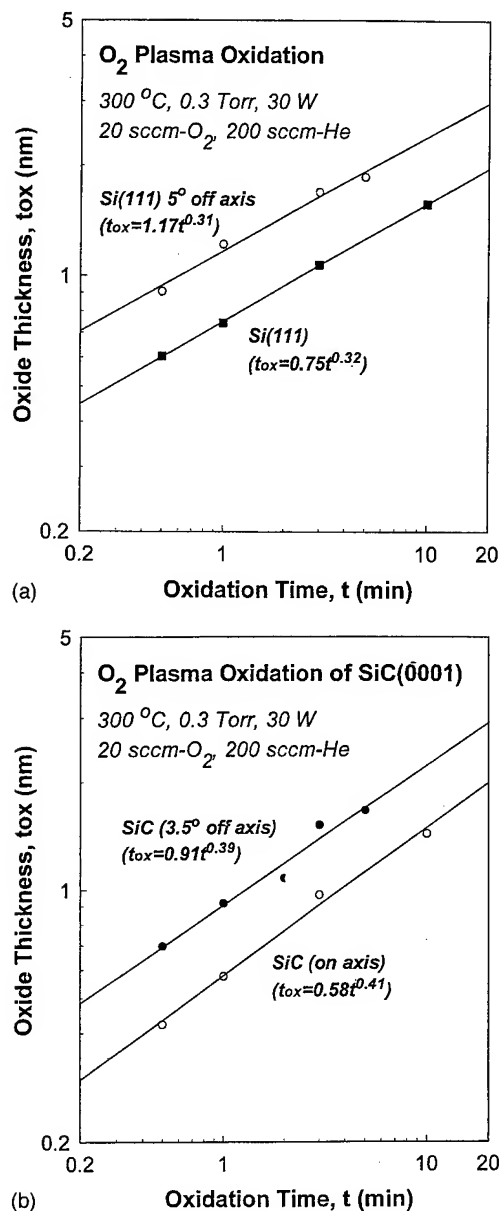


FIG. 6. Log-log plots of the oxide thickness versus the oxidation time for the O_2 RPAO of flat and vicinal (a) Si and (b) SiC. The data are fit to a power law dependence, $t_{ox} = at^b$.

- (i) the solid state oxidation product is SiO_2 ;
- (ii) there are no interfacial C-O bonds; and
- (iii) oxidation proceeds initially more rapidly on vicinal surfaces than on flat surfaces, paralleling what has been found for Si.

The primary solid state oxidation product for thermal oxidation of SiC is also SiO_2 ; however, there have been reports of the observation of C-O at or near the SiC-SiO₂ interface.³ This means that there are significant differences in the oxide formation chemistries by low temperature plasma assisted and high temperature thermal oxidation processes. For example, in the plasma assisted processes the oxidation species are typically long-lived molecular metastables such as O_2^+ or positive molecular ions such as O_2^+ , whereas in the thermal

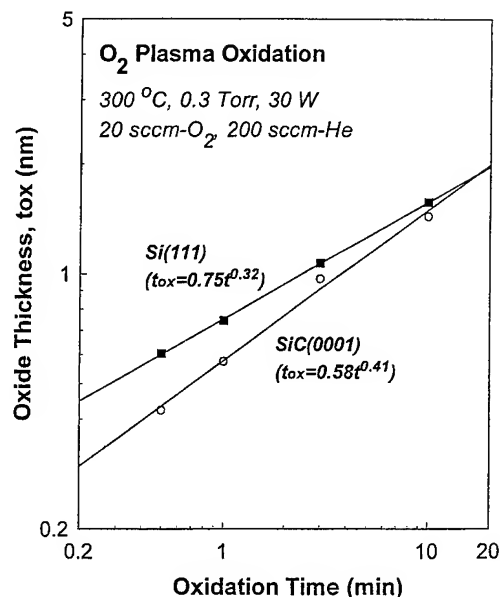


FIG. 7. Log-log plots comparing oxide thickness versus the oxidation time for the O_2 RPAO of flat Si and SiC. The data are fit to a power law dependence, $t_{ox} = at^b$.

oxidation process the oxidation species are typically O atoms. Since the thermal and plasma oxidation processes are qualitatively different with respect to the formation of C-O bonds, then the differences between the two types oxidation processes could be expected to require different post-oxidation and post-deposition procedures for forming low defect density SiC-SiO₂ interfaces. However, this is not the case, as is discussed in Sec. IV B of the article. Before dis-

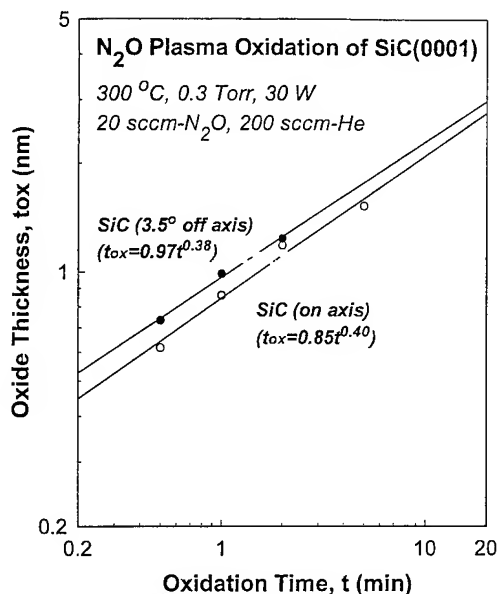


FIG. 8. Log-log plots comparing the oxide thickness versus the oxidation time for the RPAO of flat SiC using O_2 and N_2O as the oxygen atom source cases. The data are fit to a power law dependence, $t_{ox} = at^b$.

cusssing this aspect of the experimental results, additional aspects of the power laws fits are discussed below.

The power laws fits to RPAO data show that the oxidation process proceeds very rapidly initially and then slows down considerably with increasing oxidation time. This means chemical reaction rates between excited oxygen species and Si- and Si- and C atom dangling bonds on Si and SiC, respectively, are decreased after Si-O bonds and a superficial layer of SiO₂ are formed at the onset of RPAO.¹⁴

For the case of RPAO of Si the oxidation rate for flat surfaces is faster using O₂ than N₂O for the same relative flow rates of He and O₂ or N₂O. For the RPAO of SiC the situation is reversed for flat surfaces and the oxidation rate is faster using N₂O. Studies are currently underway to address this difference. A model has been proposed for RPAO of Si that explains the differences in the oxidation rates using O₂ or N₂O source gases.¹⁵ This model was developed primarily to explain the retention of nitrogen atoms at the Si-SiO₂ interface during oxide growth using the N₂O source gas. The surface contamination of SiC by nitrogen makes it impossible to determine whether there is nitrogen retention at the SiC-SiO₂ interface during oxide growth in N₂O. This issue is presently under study, and will require the development of procedures for *ex situ* surface preparation that remove residual nitrogen atom contamination from the SiC surface.

B. Electrical properties of SiC-SiO₂ interfaces

This section summarizes results of recent electrical measurements made on MOS capacitors, in particular it compares values of D_{it} as extracted from capacitance-voltage measurements on SiC interfaces prepared on *p*-type substrates. The experimental studies of Stein von Kaminski *et al.*¹¹ have demonstrated mid-gap D_{it} values in the low $10^{11} \text{ cm}^{-2} \text{ eV}^{-1}$ range for interfaces prepared by high temperature thermal oxidation. These device structures used metal electrodes. The attainment of low defect densities required two annealing steps:

- (i) a post-oxidation an Ar/H₂ mixture at 1150 °C, and
- (ii) a conventional PMA in an H₂ containing ambient at 400 °C after metallization.

When annealing procedures that did not include H₂ in the 1150 °C anneal were applied to the RPAO formed interfaces, the D_{it} values were significantly higher, in the $10^{13} \text{ cm}^{-2} \text{ eV}^{-1}$ range.⁷ However as shown in Ref. 8, these D_{it} values were significantly reduced for the RPAO interfaces following a high temperature (1150 °C) anneal in an Ar/H₂ mixture. It is interesting to note that the difference in D_{it} values between RPAO devices subjected to the Ar and Ar/H₂ anneals is very nearly equal to the density of carbon atom dangling bonds at the step edges. This suggests the possibility that carbon atom dangling bonds are not terminated by during the oxidation or oxide deposition steps, or during the conventional PMA.

Consider first the RPAO interfaces. Since our studies show no C-O bonding in interfacial regions for either the flat or vicinal wafers, it is suggested that the step edge carbon

atom dangling bonds may remain unterminated after the oxidation process. The high defect density after the conventional 400 °C PMA in a hydrogen containing ambient further suggests that the step edge carbon atom dangling bonds are not hydrogen terminated by this process. If this is the case then hydrogen atom production during the 1150 °C anneal in Ar/H₂ is sufficiently high to produce C-H bonding at the step edges. Since thermal oxidation processes show evidence for C-O bonding in interfacial transition regions³ and the RPAO process does not, this suggests that there might be differences related to the way step carbon atom dangling bonds are terminated, either during thermal and plasma oxidations, or in post-oxidation annealing. However, this appears not to be the case, since interfaces produced by high temperature thermal oxidation also require a high temperature anneal in an H₂ containing ambient. This means that even though C-O bonds can be found in silicon oxycarbide transition regions after high temperature thermal oxidation³ they may not be formed in sufficient numbers at carbon atom step edges to neutralize dangling bond defects. Alternatively, the local bonding environment of carbon atoms in interfacial regions and at the Si-C step edges is different and may be a contributing factor to oxygen atom termination.

Finally, it is also interesting to note that fabrication of low D_{it} interfaces on *n*-type SiC by RPAO and thermal oxidation does not require a high temperature anneal in an H₂ containing ambient.⁷ There are two possible explanations for not requiring such an anneal for the RPAO formed interfaces:

- (i) the active interfacial defects are in the bottom half of the SiC band gap and hence are more active in *p*-type material, or
- (ii) the formation of C-O bonds at an SiC interface is Fermi level dependent favoring C-O bond formation in *n*-type material.

The wafers that we studied were nitrogen doped *n*-type wafers, and there was no evidence of interfacial C-O bond formation. However, the concentration of step edge C atom danglings bonds that are available for termination by C-O bonding is below the AES detection limit. 6H SiC flat and vicinal SiC wafers with C faces have been ordered from Cree Research, Inc., upon their receipt, an AES study of the initial stages of PRAO will be initiated. It is clear that these questions identified with respect to electrical results on SiC devices need additional experimental and theoretical studies before they can be satisfactorily resolved. In particular, electron spin resonance experiments that can readily distinguish between Si and C atom dangling bonds may be helpful in resolving the issues discussed in the last two paragraphs.

ACKNOWLEDGMENTS

This research has been supported by the the National Science Foundation, Office of Naval Research, and Air Force Office of Scientific Research (Multi-University Research Initiative).

- ¹L. A. Lipkin and J. W. Palmour, *J. Electron. Mater.* **25**, 909 (1996).
- ²T. Ouisse, *Philos. Mag. B* **73**, 325 (1996).
- ³B. Hornetz, H.-J. Michel, and J. Halbritter, *J. Mater. Res.* **9**, 3088 (1994).
- ⁴J. T. Fitch, E. Kobeda, G. Lucovsky, and E. A. Irene, *J. Vac. Sci. Technol. B* **7**, 153 (1989).
- ⁵T. Yasuda, Y. Ma, S. Habermehl, and G. Lucovsky, *Appl. Phys. Lett.* **60**, 434 (1992).
- ⁶G. Lucovsky, Yi Ma, S. V. Hattangady, D. R. Lee, Z. Lu, V. Misra, J. J. Wortman, and J. L. Whitten, *Jpn. J. Appl. Phys.* **1** **33**, 7061 (1994).
- ⁷G. Lucovsky, A. Banerjee, B. Hinds, B. Clafin, K. Koh, and H. Yang, *J. Vac. Sci. Technol. B*, these proceedings.
- ⁸A. Gözl, R. Janssen, E. Stein von Kamienski, and H. Kurz, in *The Physics and Chemistry of SiO₂ and the Si-SiO₂ Interface*, edited by H. Z. Massoud, E. H. Poindexter, and C. R. Helms (Electrochemical Society, Pennington, NJ, 1996), p. 753.
- ⁹C. H. Bjorkman, C. E. Shearon, Jr., Y. Ma, T. Yasuda, G. Lucovsky, U. Emmerichs, C. Meyer, K. Leo, and H. Kurz, *J. Vac. Sci. Technol. B* **11**, 964 (1993).
- ¹⁰T. Yasuda, D. R. Lee, C. H. Bjorkman, Y. Ma, U. Emmerichs, C. Meyer, K. Leo, and H. Kurz, *Mater. Res. Soc. Symp. Proc.* **315**, 375 (1993).
- ¹¹E. Stein von Kamienski *et al.*, *Microelectron. Eng.* **28**, 201 (1995).
- ¹²G. Lucovsky, H. Niimi, K. Koh, D. R. Lee, and Z. Jing, in Ref. 8, p. 441.
- ¹³D. R. Wolters and A. T. A. Zegers-van Duynhoven, *J. Appl. Phys.* **65**, 5126 (1989).
- ¹⁴P. Thanikasalam, T. K. Whidden, and D. K. Ferry, *J. Vac. Sci. Technol. B* **14**, 2840 (1996).
- ¹⁵K. Koh, H. Niimi, and G. Lucovsky, *Mater. Res. Soc. Symp. Proc.* (in press).

Raman spectra as a measure of interface alloying for IV/IV superlattices

Jian Zi,^{a)} Kaiming Zhang, and Xide Xie

Surface Physics Laboratory and Fudan-T. D. Lee Physics Laboratory, Fudan University, Shanghai 200433, People's Republic of China

(Received 21 January 1997; accepted 5 May 1997)

Raman spectra of strained (001)-oriented Si/Ge and α -Sn/Ge superlattices with alloyed interfaces are calculated by a bond-polarizability model with lattice dynamical properties described by using a Keating model. The alloyed interface layers are treated by using a supercell technique. It is found that the Raman peaks around 410 cm^{-1} of Si/Ge superlattices and 260 cm^{-1} of α -Sn/Ge superlattices could be used as a measure of interface alloying. © 1997 American Vacuum Society. [S0734-211X(97)12204-1]

I. INTRODUCTION

Si/Ge superlattices (SLs) have in recent years received considerable research interest because of their potential applications in devices. Due to its excellent mechanic and etching properties, Si has become the indispensable material in semiconductor technology. Being an indirect band-gap semiconductor, however, Si has been excluded from applications such as in photonic and optoelectronic devices for years. There has been hope of improving the electronic band structure of Si in order to obtain a semiconductor material with a direct or quasi-direct band gap.¹ Si/Ge SLs have been one of the most promising candidates.^{2,3} Owing to the great advances in growth techniques, short-period α -Sn/Ge SLs have been successfully grown pseudomorphically on a Ge(001) substrate despite the very large lattice mismatch of 14.7% between α -Sn and Ge.⁴⁻⁶ The combination of a zero band-gap semiconductor and an indirect band-gap semiconductor may produce a material with novel electronic band structures and optical properties.

There has been much interest in characterizing the interfaces in the atomic scale in order to control the growth of semiconductor SLs with high-quality interfaces. The effects of imperfect interfaces have attracted considerable attention due to the fact that the interfaces of grown samples are usually imperfect. For example, owing to intermixing or interdiffusion, the interfaces are not abrupt. Alloyed interface layers may exist. Therefore, it is of importance to study the effects of alloyed interfaces not only from a purely theoretical point of view but also from a practical point of view. The effects of alloyed interfaces on the electronic and optical properties of Si/Ge SLs were studied theoretically by Theodorou and Tserbak.⁷

Raman spectroscopy has been proved to be a useful probe to gain insight into the local atomic arrangement in SLs. Recently, Raman spectroscopy has been used to characterize the interfaces of Si/Ge SLs.⁸⁻¹⁰ On the theoretical side, the effects of alloyed interfaces on the vibrational properties of Si/Ge SLs have drawn much attention.^{9,11-13} These theoretical efforts have been aimed at the origin of the additional

Raman peaks observed in experiments. It has been shown that the Raman peaks around 410 cm^{-1} in Si/Ge SLs originate from the alloyed interface layers.

In the present work, we will show that the additional Raman peaks can be used as a measure to determine the degree of interface alloying for both (001)-Si/Ge and α -Sn/Ge SLs. The article is organized as follows. In Section II the theoretical approaches adopted are briefly introduced and the calculated results are presented in Section III. Conclusions are given in Section IV.

II. THEORETICAL APPROACHES

A generalized Keating model^{14,15} is used in the present work to describe the lattice dynamical properties of Si/Ge and α -Sn/Ge SLs. For bulk materials, there are only two parameters in this model: one describes the first-nearest-neighbor bond-bending interactions and the other one the bond-stretching interactions extending up to the second-nearest neighbors. This two-parameter model could give a good description of the phonon-dispersion relations throughout the Brillouin zone except for the known flatness of TA branches which arise from the long-range interactions. This exception, however, does not affect our discussions since we are only interested in the optical phonons in the present work. The values of the Keating parameters for Si, Ge, and α -Sn are taken from Refs. 14 and 15. The force constants across the interfaces are simply assumed to take the average values of those of the two constituents. The effects of strain are properly taken into account in our model.

In order to calculate the Raman spectra, a bond-polarizability model¹⁶⁻¹⁸ is used. It has been found that this model can give a good description of the scattering intensity from optical modes. The polarizability of the whole system is calculated as a sum of independent contributions from each bond, based on the calculated eigenvalues and eigenvectors. Then, the Raman intensity in the $\mu\nu$ polarization for backscattering configuration is given by

$$I_{\mu\nu}(\omega) \propto \frac{n(\omega)+1}{\omega} \sum_j \delta(\omega-\omega_j) |\Delta\alpha_{\mu\nu}(j)|^2, \quad (1)$$

^{a)} Author to whom all correspondence should be addressed; Electronic mail: jzi@fudan.edu.cn or jianzi@fudan.ihep.ac.cn

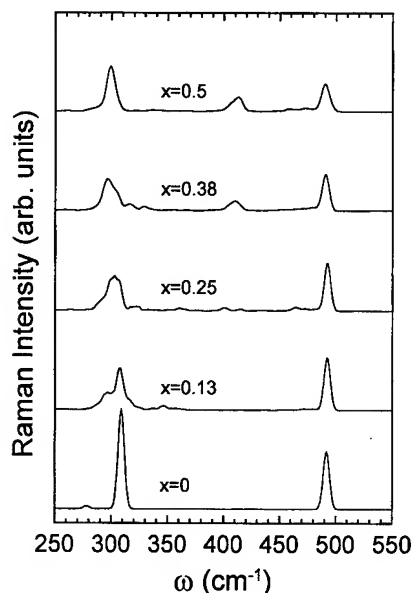


FIG. 1. Calculated Raman spectra of a Si_3/Ge_5 SL grown on a Si (001) substrate for different alloyed interfaces in the backscattering configuration $z(x,y)\bar{z}$.

where $n(\omega) + 1$ is the Bose-Einstein population factor and $\Delta\alpha_{\mu\nu}(j)$ is the variation of the polarizability tensor due to the phonon mode ω_j .

A supercell technique is used to describe the alloyed interface layers. The lattice constant of the supercell in the in-plane direction (perpendicular to the [001] growth direction) is taken to be $2a_{\parallel}$, where a_{\parallel} is the lattice constant of the substrate. Thus there are a total of eight atoms in the two-dimensional unit cell. The atomic distribution in the alloyed planes is chosen randomly for each desired alloy composition. Several distinct random configurations are used to average the final results in order to rule out some dominant local atomic structures. To simulate quantitatively the alloyed interfaces only the layers nearest to the interfaces are allowed to become alloyed. For simplicity, the degree of alloying is assumed to be identical for all alloyed layers nearest to the interfaces. For a A/B SL, the layers nearest to the interfaces in the A (B) layers become $A_{1-x}B_x$ (A_xB_{1-x}), where x stands for the degree of interface alloying.

III. CALCULATED RAMAN SPECTRA

In Fig. 1 the calculated Raman spectra of a Si_3/Ge_5 SL grown on a Si(001) substrate are shown for different interface alloying in the usually $z(x,y)\bar{z}$ backscattering configuration, in which only longitudinal modes at the zone center are Raman active.

For abrupt interface ($x=0$) two Raman peaks exist in the frequency range displayed. By inspecting the vibrational density of states (DOS) projected in each layer [see Fig. 2(a)], it is found that the Raman peak around 490 cm^{-1} originates from the first-order Si-like LO mode with vibrations sharply confined to Si layers. Due to the fact that there are only three Si layers in the SL periodicity, only a first-

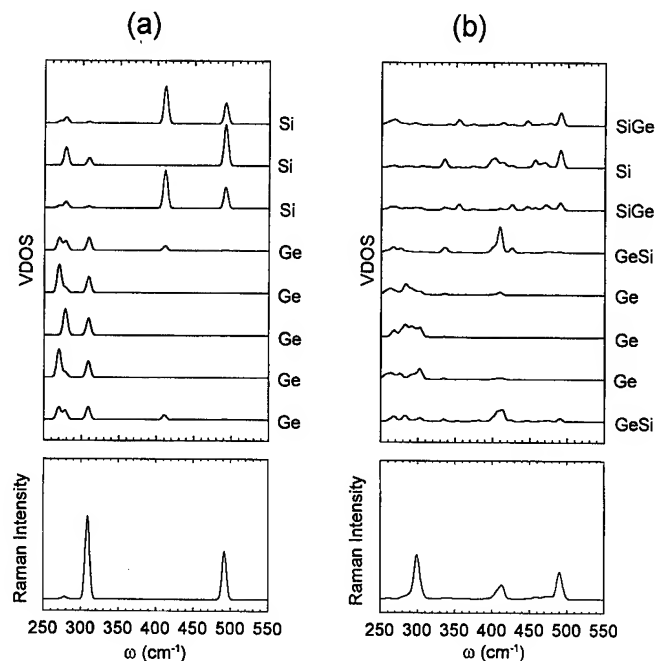


FIG. 2. Vibrational DOS projected in each layer along the [001] growth direction (upper panels) and the corresponding Raman spectra (bottom panels) of a Si_3/Ge_5 SL on a Si (001) substrate (a) with ideal interface $x=0$ and (b) with alloyed interfaces $x=0.5$.

order Si-like LO mode exists. There are three Ge-like LO modes, but only the first- and the third-order Ge-like LO modes are Raman active. The Raman peak around 310 cm^{-1} corresponds to the first-order Ge-like LO mode and the very small peak around 280 cm^{-1} to the third-order Ge-like LO mode. Due to the fact that the frequencies of Ge-like LO modes are located in the LA continuum of bulk Si, finite LA-like excitations in Si layers exist. Therefore, Ge-like LO modes are resonant, quasi-confined to Ge layers. There is a SL mode around 410 cm^{-1} , but it is Raman inactive.

With the introduction of alloyed layers at interfaces the Si-like and Ge-like peaks are broadened and an additional peak around 410 cm^{-1} appears. It becomes obvious when the degree of interface alloying is greater than 0.25. It can be seen from Fig. 1 that the intensity of this peak increases with the increase in the degree of interface alloying. In Fig. 2(b) the vibrational DOS projected at each layer is given for $x=0.5$. It is obvious the interface alloying is responsible for the 410 cm^{-1} peak observed in experiments. It is interesting to note that only the alloyed layers closest to the Ge layers contribute to the peak around 410 cm^{-1} .

In Fig. 3 the calculated Raman spectra of an $\alpha\text{-Sn}_3/\text{Ge}_5$ SL grown on a Ge(001) substrate are shown for different interface alloying in the usually $z(x,y)\bar{z}$ backscattering configuration. For an abrupt interface ($x=0$) three Raman peaks dominate the spectra in the frequency range displayed. By inspecting the vibrational DOS projected in each layer [see Fig. 4(a)] it is found that the Raman peak around 300 cm^{-1} originates from the first-order Ge-like LO mode with vibrations confined to Ge layers. The Raman peak around 255 cm^{-1} corresponds to the third-order Ge-like LO mode.

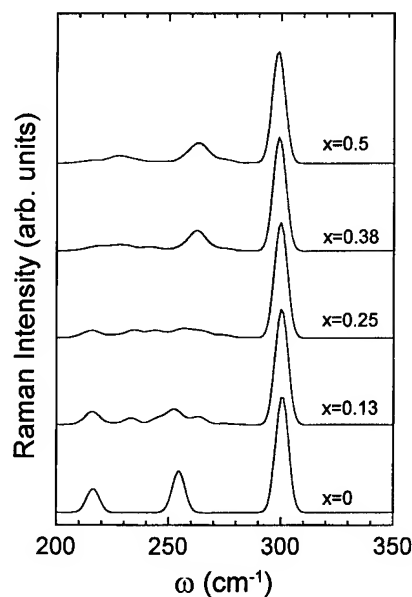


FIG. 3. Calculated Raman spectra of a α -Sn₃/Ge₅ SL grown on a Ge (001) substrate for different alloyed interfaces in the backscattering configuration $z(x,y)\bar{z}$.

There are some excitations in Sn layers for this mode due to less confinement of higher-order modes. The peak around 215 cm⁻¹ originates from the first-order Sn-like LO mode. This mode, however, is resonant, quasi-confined to Sn layers due to the fact that the LO continuum of Sn overlaps the LA continuum of Ge. As a result, finite LA-like excitations exist in Ge layers.¹⁵

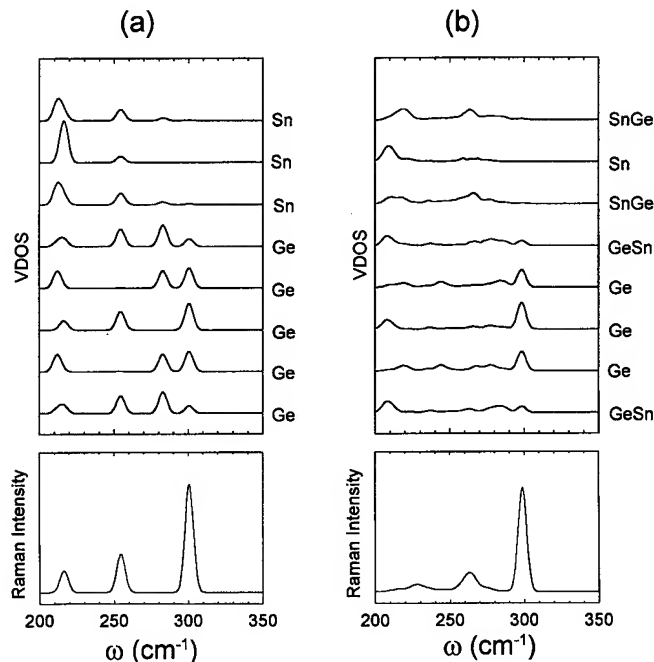


FIG. 4. Vibrational DOS projected in each layer along the [001] growth direction (upper panels) and the corresponding Raman spectra (bottom panels) of a α -Sn₃/Ge₅ SL on a Ge (001) substrate (a) with ideal interface $x=0$ and (b) with alloyed interfaces $x=0.5$.

With the introduction of alloyed layers at interfaces a new peak around 260 cm⁻¹ appears and the two peaks around 255 cm⁻¹ and 215 cm⁻¹ nearly disappear. By inspecting the vibrational DOS projected at each layer given in Fig. 2(b) for $x=0.5$, it is obvious that the peak around 260 cm⁻¹ originates from the alloyed interface layers. Similar to the case in Si/Ge SLs, only the alloyed layers closest to the Sn layers contribute to this peak. The intensity of this peak increases with the increase in the degree of interface alloying.

From our model calculations of Raman spectra for Si/Ge and α -Sn/Ge SLs we show that the intensities of additional peaks due to the alloyed interfaces increase with the increase in the degree of interface alloying. Therefore, these additional peaks can be used to characterize the degree of interface alloying of Si/Ge and α -Sn/Ge SLs from the experimentally measured Raman spectra.

IV. CONCLUSIONS

Raman spectra of Si/Ge and α -Sn/Ge SLs with alloyed interface layers were studied by using a supercell technique. It is found that with the introduction of alloyed interface layers additional Raman peaks around 410 cm⁻¹ for Si/Ge and 260 cm⁻¹ for α -Sn/Ge SLs appear. The intensities of these peaks increase with the increase in the degree of interface alloying. In other words, these two peaks can be used quantitatively as a measure to determine the interface alloying for Si/Ge and α -Sn/Ge SLs.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China and the Education Commission of China.

- ¹U. Gnatzmann and K. Clausecker, *Appl. Phys.* **3**, 9 (1974).
- ²J. Bevk, A. Ourmazd, L. C. Feldman, T. P. Persall, J. M. Bonar, B. A. Davidson, and J. P. Maennaerts, *Appl. Phys. Lett.* **50**, 760 (1987); K. Eberl, G. Krötz, R. Zachai, and G. Abstreiter, *J. Phys. (France) Colloq.* **5**, 329 (1987).
- ³U. Schmid, N. E. Christensen, M. Alouani, and M. Cardona, *Phys. Rev. B* **43**, 14597 (1991), and references therein.
- ⁴W. Wegscheider, K. Eberl, U. Menzinger, and G. Abstreiter, *Appl. Phys. Lett.* **57**, 875 (1990).
- ⁵G. Abstreiter, J. Olajos, R. Schorer, P. Vogl, and W. Wegscheider, *Semi-cond. Sci. Technol.* **8**, S6 (1993).
- ⁶R. Schorer, W. Wegscheider, and G. Abstreiter, *J. Vac. Sci. Technol. B* **11**, 1069 (1993).
- ⁷G. Theodorou and C. Tserbak, *Phys. Rev. B* **51**, 3 (1995).
- ⁸J. Menéndez, A. Pinczuk, J. Bevk, and J. P. Mannaerts, *J. Vac. Sci. Technol. B* **6**, 1306 (1988).
- ⁹M. I. Alonso, F. Cerdeira, D. Niles, M. Cardona, E. Kasper, and H. Kibbel, *J. Appl. Phys.* **66**, 5645 (1989).
- ¹⁰R. Schorer, E. Friess, K. Eberl, and G. Abstreiter, *Phys. Rev. B* **44**, 1772 (1991).
- ¹¹S. Wilk, *Superlattices Microstruct.* **9**, 267 (1992); S. Wilk, J. Maek, and B. Velicky, *Phys. Rev. B* **41**, 3769 (1990).
- ¹²J. Zi, K. Zhang, and X. Xie, *Phys. Rev. B* **47**, 9937 (1993).
- ¹³S. de Gironcoli, E. Molinari, R. Schorer, and G. Abstreiter, *Phys. Rev. B* **48**, 8959 (1993).
- ¹⁴J. Zi, K. Zhang, and X. Xie, *Phys. Rev. B* **41**, 12862 (1990).
- ¹⁵J. Zi, K. Zhang, and X. Xie, *Phys. Rev. B* **45**, 8397 (1992).
- ¹⁶A. A. Maradudin and E. Burstein, *Phys. Rev.* **164**, 1081 (1967).
- ¹⁷S. Go, H. Bilz, and M. Cardona, *Phys. Rev. Lett.* **34**, 580 (1975).
- ¹⁸B. J. Berne and R. Pecora, *Dynamic Light Scattering* (Wiley, New York, 1976).

Measurement of band offsets in Si/Si_{1-x}Ge_x and Si/Si_{1-x-y}Ge_xC_y heterojunctions

B. L. Stein and E. T. Yu^{a)}

Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, California 92093-0407

E. T. Croke and A. T. Hunter

Hughes Research Laboratories, Malibu, California 90265

T. Laursen, A. E. Bair, and J. W. Mayer

Center for Solid State Science, Arizona State University, Tempe, Arizona 85287-1704

C. C. Ahn

California Institute of Technology, Pasadena, California 91125

(Received 12 January 1997; accepted 3 April 1997)

Realization of group IV heterostructure devices requires the accurate measurement of the energy band offsets in Si/Si_{1-x}Ge_x and Si/Si_{1-x-y}Ge_xC_y heterojunctions. Using admittance spectroscopy, we have measured valence-band offsets in Si/Si_{1-x}Ge_x heterostructures and conduction-band and valence-band offsets in Si/Si_{1-x-y}Ge_xC_y heterostructures grown by solid-source molecular-beam epitaxy. Measured Si/Si_{1-x}Ge_x valence-band offsets were in excellent agreement with previously reported values. For Si/Si_{1-x-y}Ge_xC_y our measurements yielded a conduction-band offset of 100 ± 11 meV for a *n*-type Si/Si_{0.82}Ge_{0.169}C_{0.011} heterojunction and valence-band offsets of 118 ± 12 meV for a *p*-type Si/Si_{0.79}Ge_{0.206}C_{0.004} heterojunction and 223 ± 20 meV for a *p*-type Si/Si_{0.595}Ge_{0.394}C_{0.011} heterojunction. Comparison of our measured band offsets with previously reported measurements of energy band gaps in Si_{1-x-y}Ge_xC_y and Si_{1-y}C_y alloy layers indicates that the band alignment is type I for the compositions we have studied and that our measured band offsets are in quantitative agreement with these previously reported results. © 1997 American Vacuum Society. [S0734-211X(97)07204-1]

I. INTRODUCTION

Extensive research on Si/Si_{1-x}Ge_x heterostructure materials and devices has yielded impressive gains in the functionality and the performance of Si-based electronic and optoelectronic devices. However, the 4.18% lattice mismatch between Si and Ge imposes significant restrictions on composition and layer thickness in Si/Si_{1-x}Ge_x heterostructures. Recently, considerable progress has been made in the growth and characterization of Si_{1-x-y}Ge_xC_y alloys.¹⁻⁵ Si_{1-x-y}Ge_xC_y offers considerably greater flexibility, compared to that available in the Si/Si_{1-x}Ge_x material system, to control strain and electronic properties in group IV heterostructures. In particular, the smaller C atom compensates for the compressive strain present in Si_{1-x}Ge_x, leading to the possibility of fabricating group IV heterostructure devices lattice matched to Si.¹⁻⁶ Recent measurements of the total energy band gap for Si_{1-x-y}Ge_xC_y compressively strained to Si(001) indicate that incorporation of C into Si_{1-x-y}Ge_xC_y increases the band gap by 21–26 meV/%C.⁷⁻¹⁰ Effective design, fabrication, and characterization of such devices, however, additionally requires the accurate measurement of the energy band offsets in Si/Si_{1-x-y}Ge_xC_y heterojunctions.

II. EXPERIMENT

We have used admittance spectroscopy to measure both conduction-band and valence-band offsets, ΔE_c and ΔE_v ,

respectively, in Si/Si_{1-x}Ge_x and Si/Si_{1-x-y}Ge_xC_y heterojunctions. Multiple quantum well (MQW) samples consisting of 150–250 Å Si_{1-x}Ge_x or Si_{1-x-y}Ge_xC_y alternating with 350 Å Si for ten periods with dopant concentrations of approximately 7.4×10^{16} – 1×10^{17} cm⁻³, respectively, were grown by solid-source molecular-beam epitaxy on Si(100) conducting substrates. *N*-type (Sb doped) and *p*-type (B doped) structures were used for measurement of, respectively, ΔE_c and ΔE_v . These heterostructures were grown at 450 °C on 2000 Å Si buffer layers, and the *p*-type Si/Si_{0.79}Ge_{0.206}C_{0.004} heterostructure was capped by an additional 2000 Å Si layer. In addition, the *p*-type Si/Si_{0.595}Ge_{0.394}C_{0.011} heterostructure and the *n*-type Si/Si_{0.82}Ge_{0.169}C_{0.011} heterostructure were grown using Sb as a surfactant to improve structural quality.¹¹ In all cases, the thickness of the MQW structure was below the critical thickness for strain relaxation.¹² X-ray diffraction (XRD), ion channeling, and transmission electron microscopy (TEM) were performed on these samples to confirm their high structural quality. The Ge concentration was determined using Rutherford backscattering, and the C concentration was then determined by applying a strain compensation ratio for Ge:C of 9.44:1, which is given by a linear interpolation of lattice constants between Si, Ge, and SiC, to the XRD patterns. In addition, secondary ion mass spectroscopy (SIMS) was used to determine the total C concentration. Schottky barrier diodes required for the admittance measurements were formed by deposition of Cr/Au circular contacts 150–300 μm in

^{a)} Author to whom correspondence should be addressed.

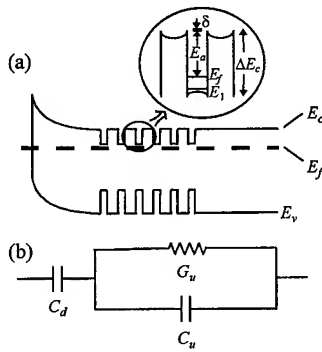


FIG. 1. (a) Energy band diagram for a Schottky barrier on an *n*-type multiple quantum well (MQW) heterostructure. The first confined state (E_1), the Fermi level relative to E_1 (E_f), and the barrier lowering due to tunneling (δ) are added to the activation energy (E_a) to obtain the band offset (ΔE_c). (b) Equivalent circuit model for this structure, including the depletion layer capacitance (C_d), and the capacitance and conductance of the undepleted portion of the sample (C_u and G_u , respectively).

diameter, followed by a mesa etch in a CF₄O₂ plasma. Al Ohmic contacts were then deposited on the backsides of the samples.

Admittance spectroscopy has been used to measure band offsets in a variety of material systems.¹³⁻¹⁵ Figure 1(a) shows a band diagram of a Schottky barrier on an *n*-type MQW structure. This structure can be modeled using an equivalent circuit consisting of the depletion layer capacitance, C_d , in series with the parallel capacitance and conductance of the undepleted portion of the sample, C_u and G_u , respectively, as shown in Fig. 1(b). In admittance spectroscopy, the total capacitance and conductance are measured as functions of temperature at various frequencies. It is generally assumed that the only circuit element that is temperature dependent is the conductance of the undepleted region, which can be modeled by thermionic emission over the Si barriers and can, therefore, be expressed as¹⁶

$$G_u = \frac{q^2 A v_{th}(T) N_c(T)}{2kT} \exp\left(-\frac{E_a}{kT}\right), \quad (1)$$

where q is the electron charge, A is the device area, v_{th} is the thermal velocity of the carriers, N_c is the effective density of states, k is Boltzmann's constant, T is the temperature, and E_a is the activation energy for emission over the QW barriers. The exponential temperature dependence of G_u is expected to be far stronger than any temperature dependences of the other circuit parameters. A resonance in the circuit shown in Fig. 1(b) occurs when $G_u = 2\pi f(C_u + C_d)$. For an admittance measurement at a fixed frequency f , this resonance will occur at a temperature $T = T_m$, at which a peak in the conductance and step in the capacitance as functions of temperature will be observed. The measurement frequency is related to the temperature at which resonance occurs by the expression

$$f = \alpha k T_m \exp(-E_a/kT_m), \quad (2)$$

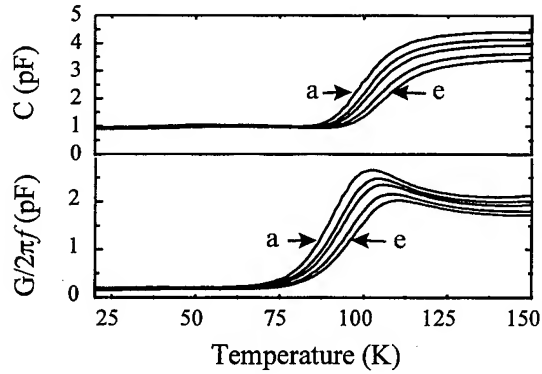


FIG. 2. Capacitance and conductance of a *p*-type Si/Si_{0.79}Ge_{0.206}C_{0.004} MQW structure as functions of temperature for (a) 1 MHz, (b) 800 kHz, (c) 600 kHz, (d) 500 kHz, and (e) 400 kHz.

where α is independent of temperature. An Arrhenius plot of $\ln(f/kT_m)$ vs $1/kT_m$ will, therefore, yield the activation energy, E_a .

As shown in the inset in Fig. 1(a), ΔE_c is related to the activation energy for an *n*-type structure by the expression¹⁶

$$\Delta E_c = E_a + E_f + E_1 + \delta, \quad (3)$$

where E_1 is the energy of the first confined state in the well, E_f is the Fermi level relative to E_1 , and δ is the barrier lowering due to tunneling. An analogous expression may be derived for the valence band. E_f is assumed to be temperature independent over the range of temperatures within which the conductance peaks occur and is calculated as the ratio of the two-dimensional carrier concentration and density of states.¹⁶ It was estimated that the error introduced by this assumption did not exceed 5 meV. E_1 is calculated by solving Schrödinger's equation and Poisson's equation self-consistently. δ has generally been found to be extremely small (<3 meV),^{14,16,17} and is, therefore, ignored in these calculations.

III. RESULTS

We first examined several *p*-type Si/Si_{1-x}Ge_x heterostructures to verify the validity of the measurement technique. Valence-band offsets measured for these Si/Si_{1-x}Ge_x structures were found to be in excellent agreement with accepted values.^{17,18} Admittance measurements were then performed on various Si/Si_{1-x-y}Ge_xC_y heterostructures. Figure 2 shows the conductance and capacitance measured as functions of temperature for various frequencies for a *p*-type Si/Si_{0.79}Ge_{0.206}C_{0.004} MQW structure; the conductance peaks and capacitance steps arising from temperature-dependent thermionic emission from the quantum wells are clearly observed on top of a background arising from leakage currents. Figure 3 [curve (a)] shows an Arrhenius plot of f and T_m , from which an activation energy of 98 ± 10 meV is obtained. The confinement energy of the lowest hole level, E_1 , was calculated to be 11 ± 1 meV, and the Fermi energy, E_f , calculated in the manner described above, was found to be 9 ± 5 meV above the first confined state. Combining these values yields a value for ΔE_v of 118 ± 12 meV. Similar measurements on the *p*-type

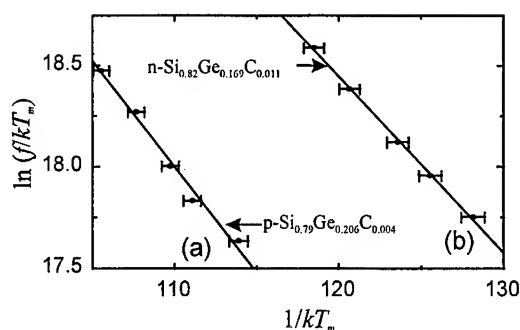


FIG. 3. Arrhenius plots of frequency (f) and the temperatures at which the conductance peaks occur (T_m) for (a) p -Si/Si_{0.79}Ge_{0.206}C_{0.004}, and (b) n -Si/Si_{0.82}Ge_{0.169}C_{0.011}.

Si/Si_{0.595}Ge_{0.394}C_{0.011} MQW structure yielded a value for ΔE_v of 223 ± 20 meV. The C concentrations determined by x-ray diffraction and SIMS for these samples were in very close agreement, indicating that the C in this sample was mostly substitutional. Information is not currently available on the presence of traps in these structures, but comparison of the admittance spectra of the Si/Si_{1-x}Ge_x samples to the Si/Si_{1-x-y}Ge_xC_y samples provides evidence that the peak observed is due to thermal activation of carriers over the barriers.

Admittance measurements were then performed on an n -type Si/Si_{0.82}Ge_{0.169}C_{0.011} MQW sample to determine ΔE_c . The C concentrations in this sample determined by x-ray diffraction and by SIMS were 1.1% and 2.7%, respectively, indicating that a significant fraction of the C was incorporated nonsubstitutionally. Although this nonsubstitutional C could influence the electronic properties of the sample, overall structural quality determined by x-ray diffraction and TEM was good, and SIMS and TEM both indicate that the C concentration is uniform throughout the alloy layers in the structure. Figure 4 shows the conductance and capacitance for this sample measured as a function of temperature at various frequencies, displaying the peaks arising from thermal activation of carriers over the barriers, with an activation energy of 91 ± 9 meV. Preliminary deep-level transient spectroscopy measurements on this structure suggest

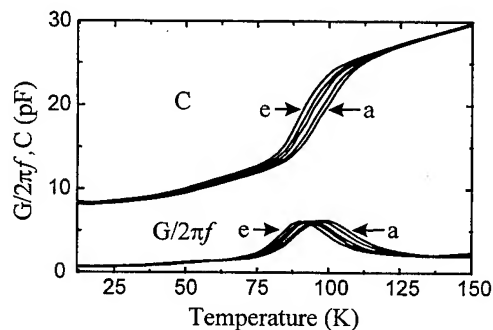


FIG. 4. Capacitance and conductance of an n -type Si/Si_{0.82}Ge_{0.169}C_{0.011} MQW structure as a function of temperature for (a) 1 MHz, (b) 800 kHz, (c) 600 kHz, (d) 500 kHz, and (e) 400 kHz.

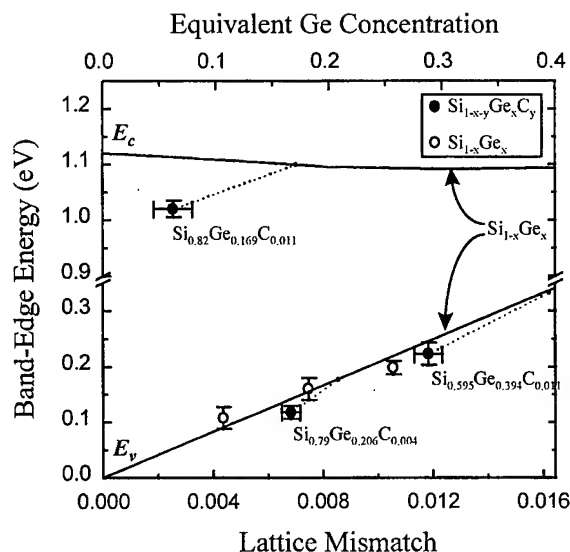


FIG. 5. Valence- and conduction-band-edge energies as a function of lattice mismatch and equivalent Ge concentration for Si_{1-x}Ge_x (open circles) and Si_{1-x-y}Ge_xC_y (closed circles) determined using our measured band offset values. The solid lines represent interpolated band-edge energies for Si_{1-x}Ge_x, and the dotted lines indicate the effect of C incorporation into a Si_{1-x-y}Ge_xC_y alloy with fixed Ge concentration.

the presence of several deep levels at least 230 meV below E_c , but in concentrations not exceeding $\sim 3 \times 10^{15} \text{ cm}^{-3}$. At measurement parameters corresponding to the resonance in the admittance spectra, no traps were observed to a sensitivity of $\sim 2 \times 10^{13} \text{ cm}^{-3}$, implying this resonance is due to thermal activation over the barriers. Using the activation energy derived from the conductance peaks, a value for ΔE_c of 100 ± 11 meV is obtained.

IV. DISCUSSION

Comparisons of our measured Si/Si_{1-x-y}Ge_xC_y band offset values with reported results for the change in the total energy band gap, ΔE_g , of 21–26 meV/%C for Si_{1-x-y}Ge_xC_y, compressively strained to Si(001),^{7–10} indicate that the band alignment for our samples is type I, and furthermore, show that our band offset values are in quantitative agreement with reported values for ΔE_g over the range of compositions for which we have measured the band offsets. Combining our measurement of ΔE_c with a value for ΔE_g of 23.5 ± 2.5 meV/%C, we would expect the values of ΔE_v for the Si/Si_{0.79}Ge_{0.206}C_{0.004} and the Si/Si_{0.595}Ge_{0.394}C_{0.011} samples to be 132 ± 11 and 219 ± 12 meV, respectively. In comparison to our measured values of 118 ± 12 and 223 ± 20 meV, respectively, these numbers are, to within the error, in very close agreement. Figure 5 shows conduction- and valence-band-edge energies for Si_{1-x}Ge_x and Si_{1-x-y}Ge_xC_y as determined from our band offset measurements. As shown in Fig. 5, incorporation of C in Si_{1-x-y}Ge_xC_y decreases both the conduction-band-edge and the valence-band-edge energies compared to those for pure Si_{1-x}Ge_x, while increasing the total energy band gap as compared to Si_{1-x}Ge_x with the same Ge concentration. The

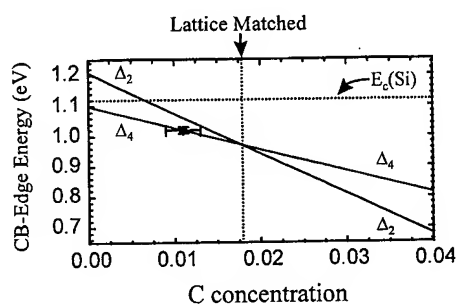


FIG. 6. Extrapolation of Δ_2 and Δ_4 conduction-band-edge energies to larger C concentrations for 16.9% Ge. Substantial conduction-band offsets may be achievable for heterostructures that are lattice matched or under tensile strain.

lowering of the band-edge energies is greater in both cases than that expected from strain compensation alone, suggesting that the compositional influence of C on the electronic structure of Si_{1-x-y}Ge_xC_y is significant.

We may extrapolate our measured band offsets to a wider range of Ge and C composition using the model-solid approach described by Van de Walle.¹⁹ Using this approach, which systematically separates the compositional or "alloy" contribution from the strain effect on the band offset, we have confirmed that our measured value of ΔE_c is in agreement with estimates of Si/Si_{1-y}C_y conduction-band offsets obtained from electrical²⁰ and photoluminescence²¹ measurements. Figure 6 shows our extrapolated conduction-band-edge energies for Si_{1-x-y}Ge_xC_y. As shown in Fig. 6, our estimates suggest that values for ΔE_c of ~200 meV or higher appear to be attainable in Si/Si_{1-x-y}Ge_xC_y heterostructures coherently strained to Si(001) for C concentrations of ~3%. Si/Si_{1-x-y}Ge_xC_y heterojunctions may, therefore, provide an attractive alternative to Si/Si_{1-x}Ge_x grown on strain-relaxed Si_{1-x}Ge_x buffer layers for fabrication of *n*-type heterostructure devices. In the valence band, extrapolations to higher C concentrations are somewhat problematic because the low C concentration for which we have measured ΔE_v tends to result in large uncertainties at higher concentrations.

V. CONCLUSION

In conclusion, we have used admittance spectroscopy to measure both ΔE_c and ΔE_v for Si/Si_{1-x}Ge_x and Si/Si_{1-x-y}Ge_xC_y heterojunctions. These measurements have shown that incorporation of C into Si_{1-x-y}Ge_xC_y lowers both the conduction- and valence-band-edge energies, while increasing the total band gap, as compared to Si_{1-x}Ge_x. This increase in the band gap is consistent with previously re-

ported values for ΔE_g in Si_{1-x-y}Ge_xC_y for the range of compositions we have measured. Moreover, our measured value for ΔE_c is in excellent agreement with previously reported ΔE_c values for Si/Si_{1-y}C_y, and in addition, suggests that significant conduction-band offsets may be achievable for Si/Si_{1-x-y}Ge_xC_y heterostructures with C concentrations of ~3%, providing a possible alternative to Si/Si_{1-x}Ge_x heterostructures grown on strain-relaxed Si_{1-x}Ge_x buffer layers for fabricating *n*-type heterostructures.

ACKNOWLEDGMENTS

The authors would like to acknowledge support from DARPA MDA972-95-3-0047 for work at UCSD, HRL, and ASU, and from ONR Grant No. N00014-95-1-0996 for work at UCSD. E. T. Y. would like to acknowledge receipt of a Sloan Research Fellowship. Two of the authors, B.L.S. and E.T.Y., would also like to thank S. S. Lau for part of the equipment used in this work.

- ¹K. Eberl, S. S. Iyer, S. Zollner, J. C. Tsang, and F. K. LeGoues, *Appl. Phys. Lett.* **60**, 3033 (1992).
- ²P. Boucaud, C. Guedj, F. H. Julien, E. Finkman, S. Bodnar, and J. L. Regolini, *Thin Solid Films* **278**, 114 (1996).
- ³H. J. Osten, H. Rücker, M. Methfessel, E. Bugiel, S. Ruminov, and G. Lippert, *J. Cryst. Growth* **157**, 405 (1995).
- ⁴J. Kolodzey, P. R. Berger, B. A. Orner, D. Hits, F. Chen, A. Khan, X. Shao, M. M. Waite, S. Ismat Shah, C. P. Swann, and K. M. Unruh, *J. Cryst. Growth* **157**, 386 (1995).
- ⁵C. W. Liu, A. St. Amour, J. C. Sturm, Y. R. J. Lacroix, M. L. W. Thewalt, C. W. Magee, and D. Eaglesham, *J. Appl. Phys.* **80**, 3043 (1996).
- ⁶R. A. Soref, *J. Appl. Phys.* **70**, 2470 (1991).
- ⁷A. St. Amour, C. W. Liu, J. C. Sturm, Y. Lacroix, and M. L. W. Thewalt, *Appl. Phys. Lett.* **67**, 3915 (1995).
- ⁸L. D. Lanzarotti, A. St. Amour, C. W. Liu, J. C. Sturm, J. K. Watanabe, and N. D. Theodore, *IEEE Electron Device Lett.* **17**, 334 (1996).
- ⁹P. Boucaud, C. Francis, F. H. Julien, J.-M. Lourtioz, D. Bouchier, S. Bodnar, B. Lambert, and J. L. Regolini, *Appl. Phys. Lett.* **64**, 875 (1994).
- ¹⁰K. Brunner, W. Winter, and K. Eberl, *Appl. Phys. Lett.* **69**, 1279 (1996).
- ¹¹P. O. Petterson, C. C. Ahn, T. C. McGill, E. T. Croke, and A. T. Hunter, *Appl. Phys. Lett.* **67**, 2530 (1995).
- ¹²R. People and J. C. Bean, *Appl. Phys. Lett.* **47**, 322 (1985).
- ¹³D. V. Lang, *Heterojunction Band Discontinuities: Physics and Device Applications*, edited by F. Capasso and G. Margaritondo (Elsevier Science, New York, 1987), Chap. 9.
- ¹⁴K. Nauka, T. I. Kamins, J. E. Turner, C. A. King, J. L. Hoyt, and J. F. Gibbons, *Appl. Phys. Lett.* **60**, 195 (1992).
- ¹⁵S. R. Smith, F. Szmulowicz, and G. J. Brown, *J. Appl. Phys.* **75**, 1010 (1994).
- ¹⁶R. E. Cavicchi, D. V. Lang, D. Gershoni, A. M. Sergent, J. M. Vandenberg, S. N. G. Chu, and M. B. Panish, *Appl. Phys. Lett.* **54**, 739 (1989).
- ¹⁷C. G. Van de Walle and R. M. Martin, *Phys. Rev. B* **34**, 5621 (1986).
- ¹⁸E. T. Yu, J. O. McCaldin, and T. C. McGill, *Solid State Phys.* **46**, 1 (1992).
- ¹⁹C. G. Van de Walle, *Phys. Rev. B* **39**, 1872 (1989).
- ²⁰W. Faschinger, S. Zerlauth, G. Bauer, and L. Palmetshofer, *Appl. Phys. Lett.* **67**, 3933 (1995).
- ²¹K. Brunner, K. Eberl, and W. Winter, *Phys. Rev. Lett.* **76**, 303 (1996).

Interfaces of strained layer $(\text{Ge}_n\text{Si}_m)_p$ superlattices studied by second-harmonic generation

Xudong Xiao,^{a)} Chun Zhang, A. B. Fedotov, Zhenghao Chen, and M. M. T. Loy
Department of Physics, The Hong Kong University of Science and Technology, Hong Kong

(Received 12 January 1997; accepted 23 March 1997)

Second-harmonic (SH) generation has been applied to study the interfaces of short-period $(\text{Ge}_n\text{Si}_m)_p$ strained layer superlattices. From the SH signal and its symmetry, we have concluded that (i) over an area of optical wavelength size, the superlattices were grown with nonuniform layer thickness in each period; (ii) along the growth direction, the atomic structure of Si-on-Ge and Ge-on-Si interfaces are not symmetric; and (iii) a structural change of the interfaces, possibly from the increase of misfit dislocation or other defect density, occurs when the thickness of the strained Ge layer in each period exceeds about six layers. The SH spectroscopic results showed defects contributed dominantly for $(\text{Ge}_7\text{Si}_8)_4$, but substantial contribution from Si-Ge bonds for $(\text{Ge}_4\text{Si}_4)_3$ and $(\text{Ge}_5\text{Si}_5)_3$ still exists. © 1997 American Vacuum Society. [S0734-211X(97)08404-7]

I. INTRODUCTION

The possibility of constructing direct or quasidirect band-gap semiconductor materials through zone folding has made the strained layer Si/Ge superlattices particularly appealing because it would allow total integration of optical devices with Si-based electronic technology.¹ It is the potential optoelectronic applications that have driven intensive studies for the strained layer Si/Ge superlattices. From theoretical^{2,3} and experimental investigations,^{4,5} it is found that both linear and nonlinear optical properties are critically dependent on the structure of the silicon/germanium interfaces. In particular, the second-order nonlinear optical property of such materials is mainly controlled by the interfaces since bulk Si and Ge are centrosymmetric and have vanishing second-order nonlinear optical susceptibility under dipole approximation. But, at the interfaces, the Si-Ge bonds are not centrosymmetric and are nonlinearly polarizable. Depending on the arrangement of these bonds at the interfaces, their contribution to second-order nonlinearity may or may not cancel out. Based on the assumption of idealized interfaces, it was predicted that there exists a large second-order nonlinear susceptibility for $(\text{Ge}_n\text{Si}_m)_p$ with odd n and m (referred to as odd-period superlattices) but a vanishing second-order nonlinear susceptibility for even n and m (referred to as even-period superlattices).³ Experimentally, such predicted nonlinear susceptibility was never observed, which was attributed to the incapability of precise control of layer thickness in superlattices in molecular beam epitaxy (MBE).⁵

Second-harmonic generation (SHG) is known to be surface and interface specific for bulk materials with centrosymmetry⁶ and, therefore, is ideal for studying the interfaces of Si/Ge superlattices. Although, unlike transmission electron microscopy (TEM), SHG lacks the spatial resolution needed for structural characterization at the atomic level, it is very relevant for understanding the optical property of the interfaces. Further, it allows us to obtain complementary in-

formation, in particular, the structural information averaged over the optical wavelength scale. Moreover, the spectroscopic capability of SHG can provide another degree of freedom to study the interfaces.

In this article, we will present our recent SHG results and extract the structural information on the interfaces. The observed finite second-order nonlinear susceptibility with a C_{4v} symmetry and a magnitude comparable to, or even one order of magnitude larger than the value predicted by theory for the ideal interfaces, enables us to probe the uniformity of Si and Ge layer thickness in each period, the interface abruptness, and interface asymmetry. By varying the thickness of the strained layer (Ge, in our case), we can also learn when the misfit dislocations or other misfit defects start to occur and how they might be spatially distributed.

II. EXPERIMENT

A series of strained layer $(\text{Ge}_n\text{Si}_m)_p$ superlattices were grown by MBE on Si(001) substrates with 3.5° miscut toward the [110] direction and with no miscut. The silicon wafers were chemically cleaned following the standard procedure before being introduced into the growth chamber and then deoxidized in ultrahigh vacuum at high temperature. A thick Si buffer layer (~2000 Å) was first deposited on the substrate at 550 °C to improve the substrates quality. Then alternating n layers of Ge and m layers of Si were grown with p repeating units followed by a protecting Si cap layer (~50 Å) at 450 °C. The growth rates of Ge and Si were 0.1 and 1.0 Å/s, respectively. The thickness of the layers was controlled by deposition time and later checked by high-resolution transmission electron microscopy (HRTEM).

The second-harmonic (SH) generation measurement of the $(\text{Ge}_n\text{Si}_m)_p$ samples was carried out by using the 1064 nm fundamental light from a Q-switched Nd:YAG laser operating at 10 Hz. The beam was incident on the samples with an incident angle of 45° and a peak power about 1 MW/cm². The SH signal at 532 nm was detected by a photomultiplier tube (PMT) with a gated integrator, with proper optical fil-

^{a)}Electronic mail: phxudong@usthk.ust.hk

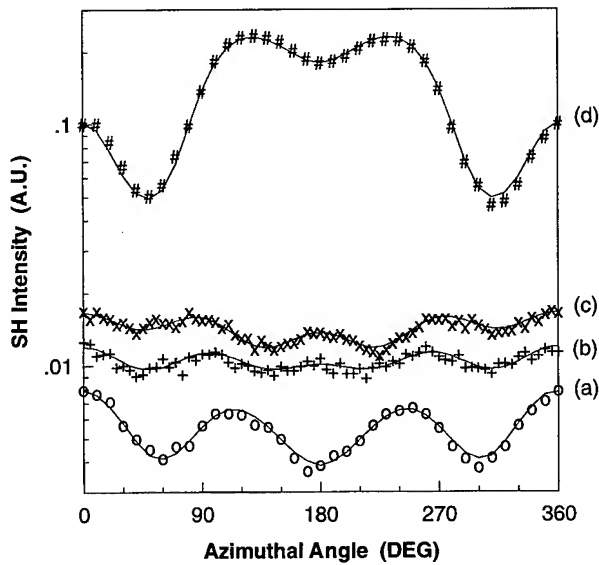


FIG. 1. SH intensity versus azimuthal angle of rotation for (a) Si(100) substrate with 3.5° miscut, (b) $(\text{Ge}_4\text{Si}_4)_3$, (c) $(\text{Ge}_5\text{Si}_5)_3$, and (d) $(\text{Ge}_7\text{Si}_8)_4$ grown on Si(100) substrates with 3.5° miscut. The incident angle is at 45°, the polarization at s -in/ p -out, and the fundamental field wavelength at 1064 nm.

tering. In order to separate the different contributions to the SHG response from the samples, the SH signal was measured as a function of sample azimuthal angle for a number of polarization combinations. In this article, we will only focus on the results with polarization combination of s input and p output. For SH spectroscopy measurement, a Ti:sapphire laser producing 100 fs pulses at 100 MHz in the wavelength range of 700–860 nm was used as the excitation source. The beam was focused to a $\sim 50 \mu\text{m}$ spot on the samples with an incident angle of 45° at an average power of 200 mW. The SH signal in the reflection was again measured by a PMT after proper filtering by color filters and a 0.25 m focal length monochromator. A lock-in technique rather than a gated integrator was employed to suppress the noise. The detection efficiency at different wavelengths was calibrated with a crystalline quartz sample for which a frequency independent $\chi_{111}^{(2)}$ could be assumed. For comparison with the

Si/Ge superlattices, bare Si(001) substrates with or without miscut were also measured.

III. RESULTS AND DISCUSSION

In Fig. 1, the measured SH signal as a function of azimuthal angle of the sample are shown for nominal $(\text{Ge}_4\text{Si}_4)_3$, $(\text{Ge}_5\text{Si}_5)_3$, and $(\text{Ge}_7\text{Si}_8)_4$ samples grown on miscut substrates together with a 3.5° miscut Si(100) substrate at a fundamental frequency of 1064 nm for s -in/ p -out polarization combination. Similar results were also observed for samples grown on substrates with no miscut. In addition to the angular independent SH signal, signal modulations with one- to fourfold rotational symmetry to various degrees were observed. All the corresponding Fourier components, obtained from fitting

$$I_{sp}^{(2\omega)}(\psi) = \left| \sum_m c_{sp}^{(m)} \cos(m\psi) \right|^2$$

to the SH signal of $(\text{Ge}_5\text{Si}_5)_3$, are listed in Table I with values of some derived second-order susceptibility components.⁷ Since the modulations with one-, three-, and fourfold symmetries are either from steps due to the substrate miscut or from the quadruples in the superlattices, they are not signatures of the interfaces. Thus, we will not discuss them further. Clearly, contributions to the SH signal from superlattices are already significant even for the two thin superlattices $(\text{Ge}_4\text{Si}_4)_3$ and $(\text{Ge}_5\text{Si}_5)_3$ samples as compared to that of the bare substrate. Noticeably, the SH signal for the thicker sample, $(\text{Ge}_7\text{Si}_8)_4$, is about one order of magnitude larger and is still dominated by the angular independent component, similarly to the thinner samples. In the following, we will proceed with our analysis, assuming Si–Ge bonds are the only contributor.

A. Nonuniform layer thickness in each period of superlattices

The twofold rotational symmetric SH signal reflects the uniformity of the superlattice layer thickness, namely the constancy of the number of layers n and m in each period of the superlattice $(\text{Ge}_n\text{Si}_m)_p$ over areas of optical wavelength size. Based on ideal uniformity (constant n for Ge and m for Si over the entire surface area), flat (no steps), and abrupt (no

TABLE I. Various Fourier components obtained by fitting the SHG data for $(\text{Ge}_5\text{Si}_5)_3$. The $\chi^{(2)}$ values were derived by a calibration with crystalline quartz. (As pointed out in Ref. 5, such fitting results by Fourier analysis are not necessarily unique. For example, for $|c_{sp}^{(2)}|$, we found the value can range from 0 to 0.018 within a tolerance of 3%.) Consider $C_{sp}^{(3)} = 1.00$ for the reference quartz.

Term	Contribution	Value of $(\text{Ge}_5\text{Si}_5)_3$	$\chi^{(2)}$ component
$C_{sp}^{(0)}$	C_{4v} symmetry of surface/interface, c_{lv} symmetry of vicinal substrate, quadrupole contribution;	0.113	$\chi_{311}^{(2)} \sim 1.0 \times 10^{-7}$ esu
$c_{sp}^{(1)}$	C_{lv} symmetry of vicinal substrate,	$0.011 + 0.047i$	$\chi_{321}^{(2)} \leq 1.0 \times 10^{-8}$ esu
$c_{sp}^{(2)}$	bulk-dipole-allow effect, C_{lv} symmetry of vicinal substrate;	$-0.010 - 0.015i$	
$c_{sp}^{(3)}$	C_{lv} symmetry of vicinal substrate;	$0.002 + 0.020i$	
$c_{sp}^{(4)}$	quadrupole contribution;	$0.000 + 0.006i$	

intermixing between Si and Ge at the atomic scale) interface assumption, the theory predicts a $\chi_{123}^{(2)}$ and $\chi_{321}^{(2)}$ and their equivalent components on the order of $\sim 1 \times 10^{-7}$ esu at 1064 nm for $(\text{Ge}_5\text{Si}_5)_3$ and 0 for $(\text{Ge}_4\text{Si}_4)_3$ under dipole approximation.³ However, the measured $\chi_{321}^{(2)}$ has an upper-bound value of $\sim 1 \times 10^{-8}$ esu for superlattice $(\text{Ge}_5\text{Si}_5)_3$, one order of magnitude smaller than the theoretical value. Not surprisingly, a similar value was also obtained for the even-period superlattices sample $(\text{Ge}_4\text{Si}_4)_3$. *The disagreement with theory can only (see discussion below) be explained by the nonuniformity of the superlattice layer thickness in each period over an area of optical wavelength size, which resulted from nonideal layer-by-layer growth.* It was recognized that even if the Si/Ge superlattice were grown via an ideal layer-by-layer growth mode (namely, the next layer would not start to grow if the present one has not been completed) on a Si(100) substrate consisting of *single-layer* steps, the $\chi_{123}^{(2)}$ and $\chi_{321}^{(2)}$ components would still vanish for the odd-period superlattices, since the two phases rotated by 90° with respect to each other in the adjacent terraces at an interface has a resultant C_{4v} symmetry after averaging over an area of optical wavelength size.⁵ However, growing Si/Ge superlattices in the ideal layer-by-layer mode on miscut Si(100) by 3.5° should recover the results from the ideal interface assumption. This is because the substrate surface is now dominated by only one single phase (in which the dimers are parallel to the steps) separated by *double-layer* steps.⁸ Such a surface regains the C_{2v} symmetry for ideal interfaces. However, experimentally, we observed vanishing $\chi_{321}^{(2)}$ for samples grown on the 3.5° miscut Si(100) substrates. This indicates that single-layer steps must have developed in the interfaces during the growth, possibly because of the Stranski–Krastanov growth mode in operation for Ge on Si(100).⁹ In another words, the samples were grown inevitably with a mixture of even and odd number of layers of Ge (or Si) simultaneously at different regions of the surface, independent of the external control on the amount of material. For example, a nominal $(\text{Ge}_5\text{Si}_5)_p$ sample would be, at a minimum, a mixture of Ge_4Si_4 , Ge_5Si_5 , and Ge_6Si_6 within the same period of superlattices at different regions [see Fig. 2(a)]. The development of single-layer steps would result in vanished $\chi_{123}^{(2)}$ and $\chi_{321}^{(2)}$ for each interface, no matter whether it is for an even- or odd-period superlattice locally. Such a growth mode makes an odd-period superlattice not purely odd period, and an even-period superlattice not purely even period. Therefore, no distinction is expected between nominal $(\text{Ge}_4\text{Si}_4)_3$ and $(\text{Ge}_5\text{Si}_5)_3$ samples as, indeed, seen in the experiment. Our conclusion here is in agreement with the TEM (Ref. 10) and x-ray diffraction (Ref. 11) and Raman spectroscopy¹² measurements, from which nonuniform layer thicknesses in each period of the superlattices were all observed with a lateral correlation about 40 nm. We shall point out that the layer nonuniformity is intrinsically determined by the thermodynamics rather than the external shutter control, as suggested by others.⁵

The above analysis is based on an abrupt interface assumption. How does intermixing [see Fig. 2(b) for models]

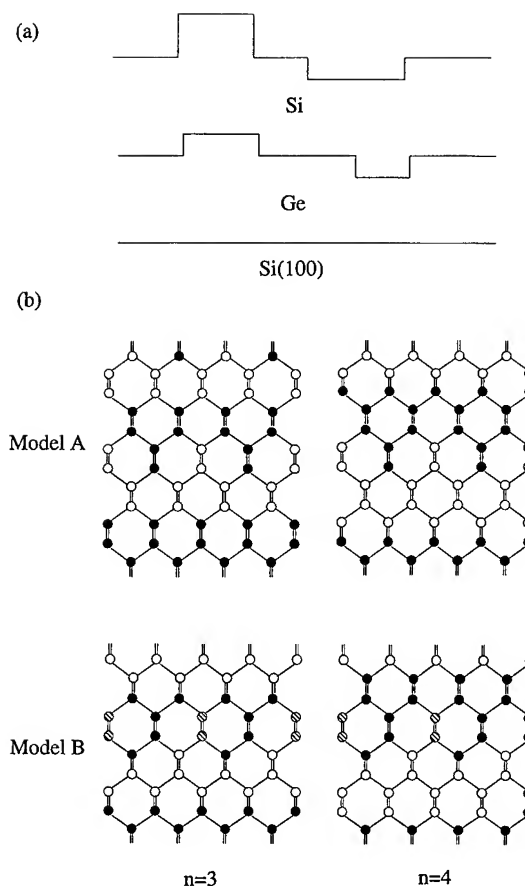


FIG. 2. (a) Sketch of the model for the nonuniform layer thickness of Ge and Si in each period of the superlattices. (b) Atomic models for intermixing of Ge and Si at the interfaces. The solid circles are Si atoms, open ones are Ge atoms, and the shaded ones are Si–Ge alloy columns.

of Ge and Si into each other at the interface affect the SH response? Various intermixing models due to Ge segregation or interdiffusion have been proposed, here, we will only analyze the two models based on the HRTEM studies.¹⁰ In Fig. 2(b), we reproduced the two models. From examining the number and orientation of the Si–Ge bonds at the two adjacent interfaces as a first-order approximation, we derive for model A, in which the Si and Ge intermix only at the Si-on-Ge interface with two layer depth and form a 2×1 structure, a zero second-order nonlinear susceptibility for odd-period superlattices (n, m are odd after averaging over a unit cell, similarly defined for even-period superlattices) but a finite $\chi_{123}^{(2)}$ and $\chi_{321}^{(2)}$ for even-period superlattices, which is *opposite* from the prediction based on the ideal abrupt interface assumption. For model B, in which Si and Ge intermix again only at the Si-on-Ge interface with a three-layer depth and form a 2×2 structure, a finite $\chi_{123}^{(2)}$ and $\chi_{321}^{(2)}$ are derived from both even- and odd-period superlattices. However, $\chi_{321}^{(2)}$ for this structure is only half the value of $\chi_{321}^{(2)}$ for the ideal interface. From the above discussion, it seems unlikely that intermixing can result in the observed vanished $\chi_{321}^{(2)}$. Single-layer steps developed in the interfaces by the nonideal layer-by-layer growth mode should be responsible for our observation.

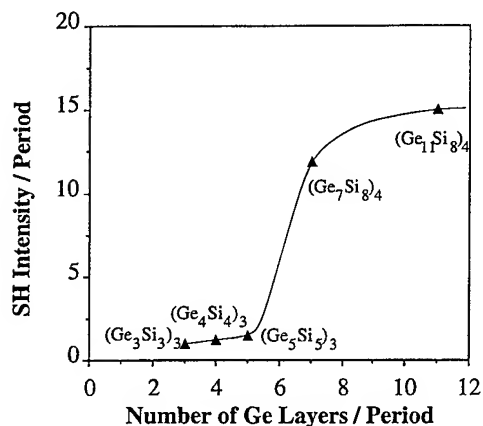


FIG. 3. The angular independent component of the SH signal for a number of Si/Ge superlattices. They are plotted versus the nominal number of Ge layers in each period.

B. Asymmetric Si-on-Ge and Ge-on-Si interfaces

The finite angular-independent SH signal is from a $\chi^{(2)}$ tensor with a C_{4v} symmetry, as verified by other polarization combinations. The measured extremely large $\chi_{311}^{(2)}$ ($\sim 1 \times 10^{-7} \text{ esu}$) component strongly indicates that the interfaces among the superlattices are not equivalent, namely the Si-on-Ge and the Ge-on-Si interfaces are not symmetric. As discussed above, any interface in the superlattice inevitably has many single-layer steps and, therefore, has a C_{4v} symmetry with a finite $\chi_{311}^{(2)}$ component. Since two adjacent interfaces, namely the Si-on-Ge and Ge-on-Si interfaces, have absolute opposite z (axis 3 for our $\chi^{(2)}$ tensor) direction, their contributions to our measurement would cancel exactly. Only if the cancellation is not perfect, can a finite $\chi_{311}^{(2)}$ be observed. This can occur if the Si-on-Ge interface is different from the Ge-on-Si interface. This conclusion holds only if a substantial amount of this SH signal is from the intrinsic interfaces. We will verify this point later. The picture reached here is consistent with the high-resolution TEM results showing the Ge-on-Si interface to be abrupt but the Si-on-Ge interface to be intermixed with an ordered structure [see Fig. 2(b) for the models].¹⁰ However, a quantitative account of the difference contributing to the nonlinear optical response between these two interfaces is not yet available.

C. SHG from dislocations and other misfit defects

In Fig. 3, we plot the angular-independent component of the SH signal averaged over the number of period p as a function of the nominal number of Ge layers n in each period. It is seen that more than a one order of magnitude increase occurs from $n=5$ to $n=7$. This large increase in SH signal indicates a structural change occurred at the interfaces once the layer thickness of the strained constituent in the superlattice exceeds certain threshold value. From other experimental studies,¹³ it was found that misfit dislocations, and in particular, misfit islands (also termed V-shaped defects⁹), started to develop at a Ge thickness above $n \sim 6$ for $(\text{Ge}_n\text{Si}_m)_p$ superlattices grown on the

Si(100) substrate in order to relieve the strain energy built in the Ge layers due to the large lattice misfit of 4.2% between Ge and Si. For the samples $(\text{Ge}_4\text{Si}_{12})_{70}$ and $(\text{Ge}_6\text{Si}_{18})_{48}$ studied by Matsuhata *et al.*, misfit islands were found to form for the latter but not for the former sample, while the number of misfit dislocations was only increased by about 30% from the former to the latter (about 1.7 dislocation/100 nm for the former and 2.2 dislocation/100 nm for the latter). By analyzing the structure of the dislocations, we find that perfect dislocations in the crystal along $[110]$, or its equivalent directions, have Si dangling bonds along them. These dangling bonds are highly nonlinear polarizable, as also evidenced from the clean Si(111) surface.¹⁴ Moreover, the network formed by these misfit dislocations has a C_{4v} symmetry when averaged over areas of optical wavelength size.¹³ For the misfit islands, the structure is more complicated.⁹ However, the five-bounded ring and seven-bounded ring units with significant distortion on the tetrahedral symmetry along the boundary of the misfit islands⁹ may also be highly nonlinear polarizable. The overall symmetry averaged over large areas for the misfit islands is also C_{4v} .¹³ Therefore, it is possible that both the misfit dislocations and misfit islands are responsible for our observed increase of the SH signal when Ge layer thickness in each period increases from $n=5$ to $n=7$. The fact that the dislocations may mainly occur at the superlattices/substrate interface¹⁵ make it even more favorable for SHG since no canceling contribution would come from the adjacent superlattice interface (see Sec. III B). The built-in asymmetry in the misfit islands (V-shaped defects) also results in noncancelled SHG between adjacent interfaces. A quantitative characterization of the misfit dislocation and misfit island densities for samples used in this study is now underway. We expect to learn the relative importance of these two types of defects in contributing to SHG.

D. SH spectroscopy and origin of SHG

If the dislocations dominate the contribution to SHG, then our earlier conclusion about the interface asymmetry (Sec. III B) can be rendered less conclusive since the misfit dislocations exist even for samples below the threshold thickness.¹³ Only if the intrinsic interfaces make a substantial contribution to SHG for $(\text{Ge}_4\text{Si}_4)_3$ and $(\text{Ge}_5\text{Si}_5)_3$, can the claim about the interface asymmetry be retained.

In Fig. 4, we show the SH spectra for two samples, $(\text{Ge}_5\text{Si}_5)_3$ and $(\text{Ge}_7\text{Si}_8)_4$. The overall features for these two spectra are similar, however, the detail of the line shapes differ significantly, as shown by the normalized spectra in the inset. Aside from the line-shape difference, the two spectra peak at about the same position, 1.67 eV, whose second-harmonic energy (2ω) coincides with the E_1 transition of silicon at ~ 3.3 eV. Such a resonance position was also observed for the Si/SiO₂ interface¹⁶ and, therefore, may have the same origin. As we discussed above, for the thicker sample $(\text{Ge}_7\text{Si}_8)_4$, the major contribution comes from the misfit defects (misfit dislocations and misfit islands), therefore, it is not surprising that its spectrum has only silicon

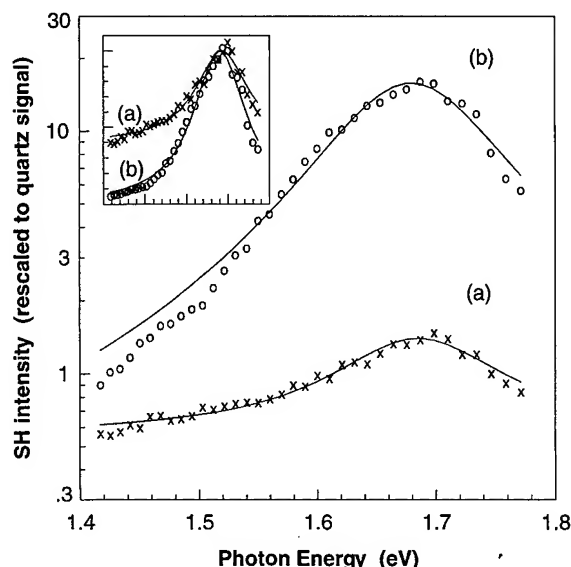


FIG. 4. SH spectra for (a) $(\text{Ge}_5\text{Si}_5)_3$ and (b) $(\text{Ge}_7\text{Si}_3)_4$ grown on Si(100) substrates. The azimuthal angles of the samples are fixed while the spectrum is taken.

characteristics. For the thin sample $(\text{Ge}_5\text{Si}_5)_3$, it is clear that defects also contributed to its SHG, as evidenced by the same resonance peak. Note, however, the significantly larger relative signal on the lower energy (longer wavelength) side of the resonance. This points to the presence of another resonance at lower energy, likely to that of the Si-Ge bond whose calculated spectrum³ has a peak at 1.2 eV. This is consistent with our picture that our measured SHG for the thin sample $(\text{Ge}_5\text{Si}_5)_3$ has a substantial contribution from the Si-Ge bonds, and is also consistent with our earlier assumption that the observed SHG at 1064 nm is dominated by intrinsic interfaces. Further spectroscopic measurement in the 1 eV energy range is required to provide the definitive test for this point.

IV. CONCLUSION

In conclusion, we have studied short-period strained layer $(\text{Ge}_n\text{Si}_m)_p$ superlattices grown on Si(100) with or without miscut by MBE by using second-harmonic generation and its spectroscopy. Analysis of the signal symmetry allowed us to conclude that the superlattices were grown with nonuniform layer thickness in each period. The observed dominating SH signal with a C_{4v} symmetry indicated an asymmetry exists between the Si-on-Ge and Ge-on-Si interfaces. By comparing the SH response for samples below and above the threshold thickness of strained Ge layers, we also identified that misfit dislocation and misfit islands could contribute to SHG significantly. Their distribution must also be uneven along the growth direction. The separation of the contribution from the misfit dislocations and misfit islands will require further

work. Finally, our SH spectroscopic results indicated that Si-Ge bonds at the interfaces could also contribute to SHG, in particular, at a high percentage for samples with a strained Ge layer thickness below the threshold value.

ACKNOWLEDGMENTS

The authors are grateful to Professor Junming Zhou of the Institute of Physics, Chinese Academy of Sciences, for his continuing interest in this work, and for his help in providing the authors with appropriate samples. This work was supported by the William Mong Solid State Clusters Laboratory and the Research Grants Council of Hong Kong.

¹See *Strained-layer Superlattices: Materials Science and Technology*, edited by T. P. Pearsall (Academic, New York, 1991).

²M. Ikeda, K. Terakura, and T. Oguchi, *Phys. Rev. B* **45**, 1496 (1991); C. Tserbak and G. Theodorou, *J. Appl. Phys.* **76**, 1062 (1994); B. A. Adderley, R. J. Turton, and M. Jaros, *Phys. Rev. B* **49**, 16 622 (1993); G. Theodorou and C. Tserbak, *ibid.* **51**, 4723 (1994).

³Ed Ghahranmani, D. J. Moss, and J. E. Sipe, *Phys. Rev. Lett.* **64**, 2815 (1990); *Phys. Rev. B* **43**, 8990 (1991); Ed Ghahranmani and J. E. Sipe, *Appl. Phys. Lett.* **62**, 2245 (1993).

⁴J. Olajos, J. Engvall, H. G. Grimmeiss, U. Menczgar, G. Abstreiter, H. Kibbel, E. Kasper, and H. Presting, *Phys. Rev. B* **46**, 12 857 (1992); T. P. Pearsall, *Appl. Phys. Lett.* **60**, 1712 (1992); M. Casalboni, N. Pinto, B. Izzi, I. Davoli, M. De Crescenzi, F. De Matteis, P. Proposito, and R. Pizzoferrato, *Phys. Rev. B* **53**, 1030 (1996); X. Zhu, Q. Xiang, M. Chu, and K. L. Wang, *J. Cryst. Growth* **150**, 1045 (1995); T. P. Pearsall, H. Polatoglou, H. Presting, and E. Kasper, *Phys. Rev. B* **54**, 1545 (1996).

⁵D. J. Bottomley, G. Lüpke, M. L. Ledgerwood, X. Q. Zhou, and H. M. van Driel, *Appl. Phys. Lett.* **63**, 2324 (1993); D. J. Bottomley, J.-M. Baribeau, and H. M. van Driel, *Phys. Rev. B* **50**, 8564 (1994).

⁶Y. R. Shen, *Annu. Rev. Phys. Chem.* **40**, 327 (1989).

⁷Chun Zhang et al. (unpublished).

⁸O. L. Alerhand, A. Nihat Berker, J. D. Joannopoulos, D. Vanderbilt, R. J. Hamers, and J. E. Demuth, *Phys. Rev. Lett.* **64**, 2406 (1990); T. W. Poon, S. Yip, P. S. Ho, and F. F. Abraham, *ibid.* **65**, 2161 (1990); E. Pehlke and J. Tersoff, *ibid.* **67**, 1290 (1991).

⁹F. K. LeGoues, M. Copel, and R. M. Tromp, *Phys. Rev. B* **42**, 11 690 (1990).

¹⁰D. E. Jesson, S. J. Pennycook, and J.-M. Baribeau, *Phys. Rev. Lett.* **66**, 750 (1991); N. Ikarashi, K. Akimoto, T. Tatsumi, and K. Ishida, *ibid.* **72**, 3189 (1994).

¹¹R. L. Headrick, J.-M. Baribeau, *Phys. Rev. B* **48**, 9174 (1993); *J. Vac. Sci. Technol. B* **11**, 1514 (1993); R. L. Headrick, J.-M. Baribeau, and Y. E. Strausser, *Appl. Phys. Lett.* **66**, 96 (1995); J.-M. Baribeau, D. J. Lockwood, and R. W. G. Syme, *J. Appl. Phys.* **80**, 1450 (1996); E. Koppensteiner, P. Hamberger, G. Bauer, V. Holy, and E. Kasper, *Appl. Phys. Lett.* **64**, 172 (1994); T. Tamagawa, T. Shintani, H. Ueba, C. Tatsuyama, K. Nakagawa, and M. Miyao, *Thin Solid Films* **237**, 282 (1994); P. M. Adams, R. C. Bowman, Jr., C. C. Ahn, S. J. Chang, V. Arbet-Engels, M. A. Kallel, and K. L. Wang, *J. Appl. Phys.* **71**, 4305 (1992).

¹²K. Dettmer, W. Freiman, M. Levy, Yu. L. Khait, and R. Beserman, *Appl. Phys. Lett.* **66**, 2376 (1995); O. Brafman, R. Manor, M. A. Aroujo Silva, F. Cerdeira, and J. C. Bean, *Physica B* **219&220**, 502 (1996).

¹³H. Matsuhata, K. Miki, K. Sakamoto, T. Sakamoto, and S. Yoshida, *Phys. Rev. B* **47**, 10 474 (1993).

¹⁴W. Daum, H.-J. Krause, U. Reichel, and H. Ibach, *Phys. Rev. Lett.* **71**, 1234 (1993).

¹⁵R. Hull, J. C. Bean, F. Cerdeira, A. T. Fiory, and J. M. Gibson, *Appl. Phys. Lett.* **48**, 56 (1986).

¹⁶C. Meyer, G. Lüpke, U. Emmerichs, F. Wolter, and H. Kurz, *Phys. Rev. Lett.* **74**, 3001 (1995).

Two-dimensional electron gas in AlGaIn/GaN heterostructures

J. Z. Li, J. Y. Lin, and H. X. Jiang^{a)}

Department of Physics, Kansas State University, Manhattan, Kansas 66506-2601

M. A. Khan and Q. Chen

APA Optics, Inc., Blaine, Minnesota 55449

(Received 12 January 1997; accepted 11 March 1997)

The formation of a two-dimensional electron gas (2DEG) system by an AlGaIn/GaN heterostructure has been further confirmed by measuring its electrical properties. The effect of persistent photoconductivity (PPC) has been observed and its unique features have been utilized to study the properties of 2DEG formed by the AlGaIn/GaN heterointerface. Sharp electronic transitions from the first to the second subbands in the 2DEG channel have been observed by monitoring the 2DEG carrier mobility as a function of carrier concentration through the use of PPC. These results are expected to have significant implications on field-effect transistor and high electron mobility transistor applications based on the GaN system. © 1997 American Vacuum Society. [S0734-211X(97)08504-1]

I. INTRODUCTION

GaN wide band-gap semiconductors have been recognized as technologically very important materials.¹⁻⁷ They have recently attracted considerable interest due to their applications for optical devices, which are active in the blue and ultraviolet (UV) wavelength regions, and electronic devices capable of operation at high-power levels, high temperatures, and harsh environments. Devices based on low band-gap materials, such as Si and GaAs, operate only in the red and near infrared wavelength regions and show poor tolerance to operation at high-power levels or elevated temperatures or in chemically hostile environments due to the low band gap, the uncontrolled generation of intrinsic carriers, and their low resistance to caustic chemicals. Due to their large dielectric strengths, GaN based devices can operate at much higher voltages for any dimensional configuration. Furthermore, they are virtually immune from environmental attack. One of the strongest motivations of the current research in the GaN system is its potential for fabricating high-power green/blue/UV lasers and electronic devices. High electron mobility transistors and field-effect transistors, based on AlGaIn/GaN heterostructures, hold promise for high-frequency microwave as well as for high-power and high-temperature electronic device applications and offer the advantage of high carrier mobilities due to the formation of two-dimensional electron gas (2DEG) by a heterojunction.^{8,9} However, practical operation of these devices still requires detailed material and device characterization and optimization.

In this work, the properties of a two-dimensional electron gas formed by an AlGaIn/GaN heterojunction interface have been probed by Hall measurements. We have observed sharp electronic transitions resulting from multiple subbands in the 2DEG channel by monitoring the 2DEG carrier mobility as a function of carrier concentration through the use of persistent photoconductivity (PPC). As a consequence of PPC, the de-

vice is sensitive to light and the sensitivity is associated with a persistent photoinduced increase in the 2DEG carrier mobility and density.

II. EXPERIMENT

As shown in Fig. 1, the device structure investigated in this work consisted of a 2 μm highly insulating GaN epilayer followed by a 25 nm thick unintentionally doped *n*-type GaN conducting channel, again followed by a 25 nm thick unintentionally doped *n*-type $\text{Al}_{0.1}\text{Ga}_{0.9}\text{N}$ epilayer. The structure was deposited over a basal plane sapphire substrate with an AlN buffer layer using a low-pressure metalorganic chemical vapor deposition (MOCVD) system. The Ohmic contacts were formed by soldering In spots directly onto the sample. The sample size was about 1 cm \times 1 cm and the Ohmic contacts were about 1 mm in diameter. The typical room-temperature carrier concentrations (due to N vacancies) were, respectively, 1×10^{17} and $5 \times 10^{17} \text{ cm}^{-3}$ for *n*-GaN and *n*- $\text{Al}_{0.1}\text{Ga}_{0.9}\text{N}$ epilayers grown under similar conditions.⁸ The 2DEG electron density and mobility were determined by variable-temperature Hall measurements. Illumination of the sample was achieved using a mercury lamp ($h\nu > E_g$ of GaN) or a neon lamp ($h\nu < E_g$ of GaN). Details of PPC characterization procedures were similar to those described previously for a *p*-type GaN epilayer.¹⁰

III. RESULTS AND DISCUSSIONS

The 2DEG carrier mobilities at different temperatures have been measured in the dark and under illuminated conditions, and the results are shown in Fig. 2. During the course of this investigation, we have used a mercury lamp ($h\nu > E_g$ of GaN) and a neon lamp ($h\nu < E_g$ of GaN) as excitation sources and found that both light sources gave similar results. Since high-temperature properties are very important for device applications based on these materials, we have replotted the 2DEG mobilities at elevated temperatures in the inset of Fig. 2. Several interesting and important

^{a)}Electronic mail: Jiang@Phys.ksu.edu

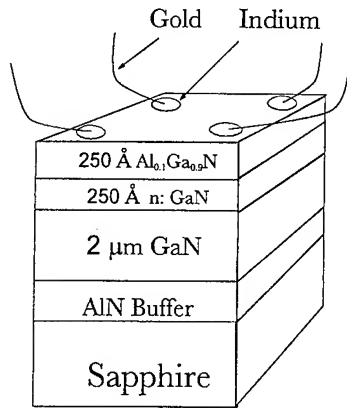


FIG. 1. Schematic diagram of the MOCVD grown AlGaIn/GaN heterostructures used in this work.

features can be observed. (i) The typical temperature dependence of the three-dimensional electron mobility of a semiconductor is absent here, which further confirms the formation of a 2DEG in the AlGaIn/GaN heterostructure. For example, in GaN epilayers, the scattering is dominated by ionized impurities at low temperatures and by phonons at higher temperatures. Hence, the electron mobility in GaN epilayers increases as temperature decreases from room temperature and reaches a maximum value between 50 and 150 K, and it then decreases as temperature further decreases due to ionized impurity scattering.^{8,11} To the contrary, the electron mobility in the 2DEG channel increases monotonically

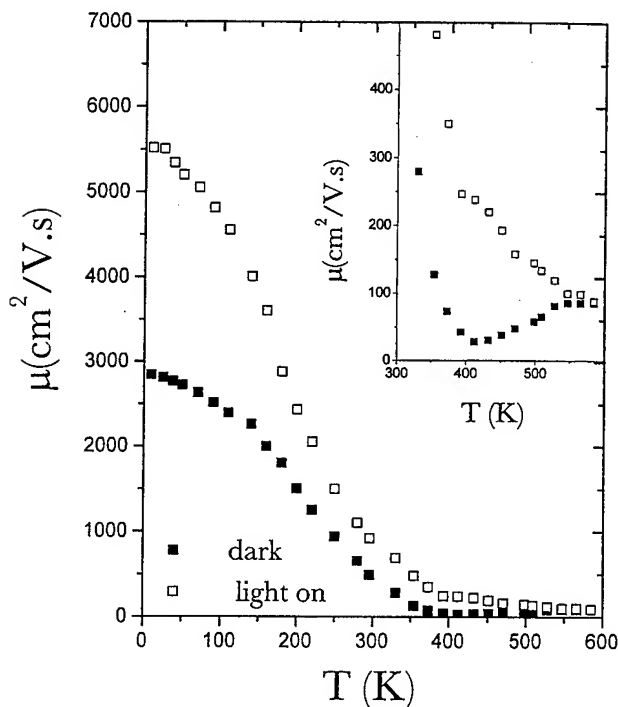


FIG. 2. 2DEG mobility, μ , as a function of temperature, T , measured in the dark (solid squares) and under illumination (open squares) conditions. For a better illustration, the 2DEG mobilities obtained at the elevated temperatures are shown in the inset.

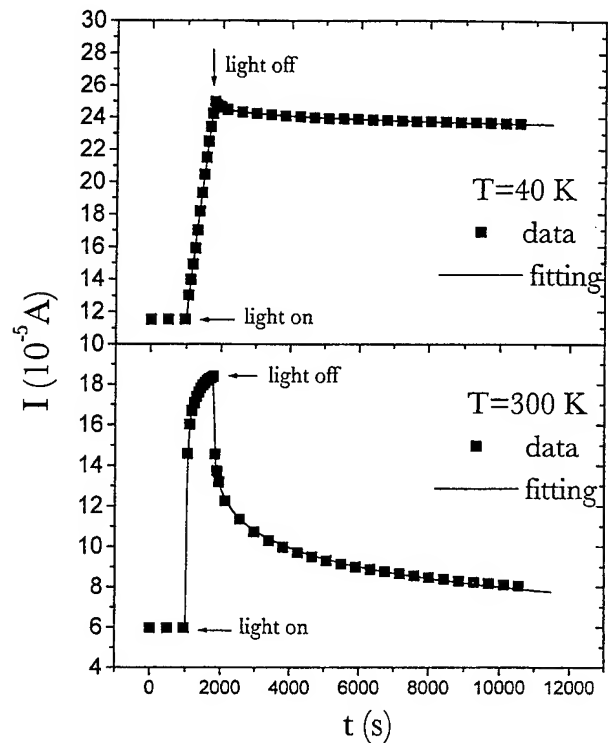


FIG. 3. PPC behavior in an AlGaIn/GaN heterostructure at different temperatures, where solid curves are the least-squares fit of the data using the formula $I_{\text{PPC}}(t) = I_d + (I_{\text{max}} - I_d)(1 - e^{-at})$ for the buildup and $I_{\text{PPC}}(t) = I_d + (I_0 - I_d)\exp[-(t/\tau)^{\beta}]$ for the decay.

with a decrease of temperature. This is a direct consequence of the ionized impurity concentration in the 2D system being almost constant due to the interface depletion effects. Hence, the monotonical decrease of the carrier mobility with temperature just reflects the fact that the electron-phonon scattering rate increases with an increase of temperature. (ii) The 2DEG mobility is enhanced significantly for all temperatures after photoexcitation. We attribute this increase to the increased electron density in the 2DEG channel, n_s , while the system is in the persistent photoconductivity state. This will be discussed further later. (iii) The 2DEG mobility in the dark state reaches a minimum value of about $25 \text{ cm}^2/\text{V s}$ at 410 K. However, the 2DEG mobility is almost one order of magnitude higher at the same temperature under illumination. Thus, it is expected that the heterojunction device performance will improve under light illumination. This is important for device applications based on GaN. (iv) The 2DEG mobilities in the dark and under illumination approach the same values at temperatures above 580 K. This is the consequence of the absence of the PPC effect at these temperatures, so the light illumination can no longer alter the carrier concentration in the 2DEG channel at high temperatures.

The conductivity in the 2DEG channel at the AlGaIn/GaN interface is extremely sensitive to light. More strikingly, as shown in Fig. 3, the photoinduced increase in the conductivity persists for a long period of time after the removal of light, an effect which is referred to as PPC. Such an effect

has been observed in AlGaAs/GaAs heterostructures only at low temperatures ($T < 150$ K).¹² As for the AlGaAs/GaAs system,^{12,13} we can identify three main mechanisms for the persistent increase in the conductivity in the 2DEG channel after illumination (or PPC) in AlGaIn/GaN heterostructures: (i) photoionization of deep-level donors in the AlGaIn barrier; (ii) photoionization of deep-level donors in the GaN layer; and (iii) the generation of electron-hole pairs in the GaN epilayer with subsequent charge separation at the interface, which requires the excitation photon energy to be larger than the energy gap of GaN. Experimentally, we found that the neon lamp ($h\nu < E_g$ of GaN) produces the PPC effect in the AlGaIn/GaN heterostructure as well, and thus, the mechanism (iii) is less likely. Furthermore, we have also performed comparison measurements on GaN epilayers grown under similar conditions and found that the PPC effect is absent in the GaN epilayers, which precludes mechanism (ii). These results then suggest that the electron density in the 2DEG channel is contributed to primarily by the transfer of photoexcited electrons from the deep-level impurities (or DX centers) in the AlGaIn epilayer.

The buildup of PPC caused by DX centers in AlGaAs has been experimentally observed and theoretically formulated to follow:¹⁴

$$I_{\text{PPC}}(t) = I_d + (I_{\text{max}} - I_d)(1 - e^{-\alpha t}), \quad (1)$$

where α is a constant, I_d is the initial dark conductivity, and I_{max} is the saturation level. While the decay of PPC associated with DX centers in AlGaAs follows a stretched-exponential function,¹⁴

$$I_{\text{PPC}} = I_d + (I_0 - I_d) \exp[-(t/\tau)^\beta], \quad \beta < 1, \quad (2)$$

where I_0 is defined as the conductivity buildup level at the moment of light excitation being terminated, τ is the PPC decay time constant, and β is the decay exponent. Figure 3 shows the buildup and decay kinetics of PPC in our AlGaIn/GaN heterostructure measured for two representative temperatures, (a) $T = 40$ K and (b) $T = 300$ K. The solid curves are the least-squares fit of data with Eq. (1) for the PPC buildup and Eq. (2) for the PPC decay. It has been demonstrated that PPC,¹⁵ when caused by the spatial separation of photogenerated electrons and holes by an electric field at a macroscopic barrier, decays logarithmically in time. Thus, our results shown in Fig. 3 also suggest that the carrier density in the 2DEG channel is most likely due to the transfer of photoexcited electrons from the deep-level impurities in the AlGaIn material. The PPC decay time constants, τ , are very long, especially at low temperatures. The fitted parameters for the PPC buildup and decay were $\alpha = 9.1 \times 10^{-4} \text{ s}^{-1}$, $\tau = 5.3 \times 10^6 \text{ s}$, and $\beta = 0.35$ at $T = 40$ K and $\alpha = 5.7 \times 10^{-3} \text{ s}^{-1}$, $\tau = 1.3 \times 10^3 \text{ s}$, and $\beta = 0.31$ at $T = 300$ K.

By utilizing the key features in the PPC state, i.e., the very long lifetimes of photoexcited charge carriers and the continuous variation of the carrier density in the 2DEG channel in a single sample, we have measured the 2DEG electron mobility, μ , as a function of the electron sheet density, n_s . Figure 4 illustrates the result for a representative temperature

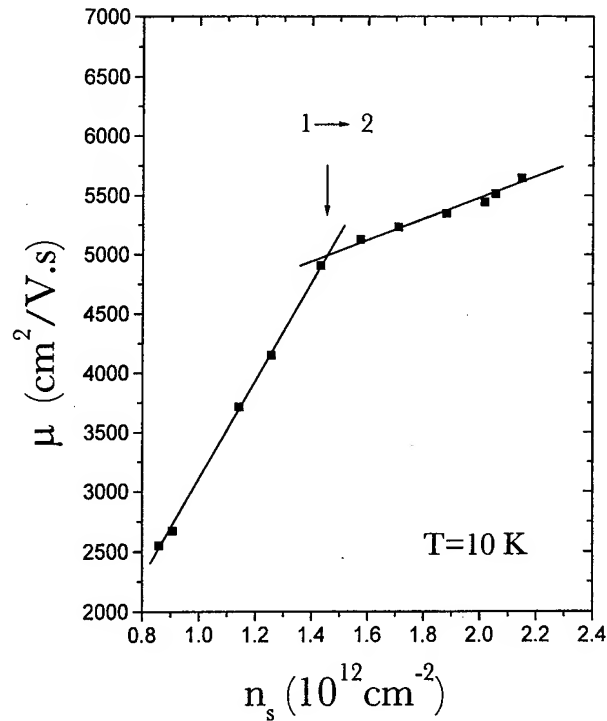


FIG. 4. The 2DEG electron mobility, μ , as a function of electron sheet density, n_s , measured at $T = 10$ K. The arrow indicates the onset electron sheet density at which the electron transport in the second subband becomes dominating.

at 10 K, which shows that μ increases almost linearly with n_s when passing from the dark to the saturated PPC state. A mobility value as high as $5800 \text{ cm}^2/\text{V.s}$ at 10 K in the PPC state is among the highest values reported for GaN. Similar trends have been observed for all temperatures up to 400 K. The mobility enhancement due to photoexcitation at a fixed temperature can be attributed to the increased electron mean energy with increasing carrier density in the 2DEG channel, which results in a less efficient interaction of the 2DEG electrons with the ionized donor impurities as well as an improved screening. Similar behaviors have been observed previously in AlGaAs/GaAs (Ref. 16) and AlInAs/GaInAs (Ref. 17) heterostructures at low temperatures. More interestingly, one should also notice that the mobility as a function of carrier density, plotted in Fig. 4, shows a sharp change in slope, as indicated by the arrow in Fig. 4. This is a clear indication of an electronic transition from the first subband to the second subband in the 2DEG channel. The carrier density at which the slope suddenly changes represents the onset density for the transition from the first to the second subband at 10 K. The total density of states of the first subband at $T = 0$ can be calculated and is about 10^{12} cm^{-2} . This value is consistent with our interpretation of the electronic transition, i.e., it is sufficiently small to allow the population of the upper subband in the PPC state.

The effects of PPC or DX centers on AlGaAs/GaAs heterojunction device characteristics have been well documented, namely sensitivity to light, a shift of the threshold

voltage, and collapse of the drain I - V characteristics.¹⁸ Thus, we expect the PPC seen here to have similar effects on the AlGaIn/GaN heterojunction device characteristics as well. Recently, an improved electrical performance of AlGaIn/GaN modulation-doped field-effect transistors due to optical excitation has been observed,⁹ which may be related to the photoinduced increase in carrier density and mobility in the 2DEG channel seen here.

In summary, electrical properties of AlGaIn/GaN heterostructures have been probed by Hall measurements. The effect of PPC has been observed and its unique features have been utilized to study the properties of 2DEG formed by AlGaIn/GaN heterojunctions. Sharp electronic transitions from the first to the second subband in the 2DEG channel have been observed by monitoring the 2DEG electron mobility as a function of electron sheet density through the use of persistent photoconductivity.

ACKNOWLEDGMENTS

The research at Kansas State University is supported by ARO (DAAH04-96-1-0371), BMDO/ONR (N00014-96-1-0885), and DOE (96ER45604/A000).

¹H. Morkoc, S. Strite, G. B. Gao, M. E. Lin, B. Sverdlov, and M. Burns, J. Appl. Phys. **76**, 1363 (1994).

²S. N. Mohammad, A. Salvador, and H. Morkoc, Proc. IEEE **83**, 1306 (1995).

³M. Asif Khan, M. S. Shur, J. N. Kuznia, Q. Chen, J. Burn, and W. Schaff, Appl. Phys. Lett. **66**, 1083 (1995).

⁴J. I. Pankove, Mater. Res. Soc. Symp. Proc. **97**, 409 (1987).

⁵S. Nakamura, T. Mukai, and M. Senoh, Appl. Phys. Lett. **64**, 1687 (1994).

⁶N. Koide, H. Kato, M. Sassa, S. Yamasaki, K. Manabe, M. Hashimoto, H. Amano, K. Hiramatsu, and I. Akasaki, J. Cryst. Growth **115**, 639 (1991).

⁷H. Morkoc, Mater. Sci. Eng. B **43**, 137 (1997).

⁸M. Asif Khan, Q. Chen, C. J. Sun, J. W. Yang, M. S. Shur, and H. Park, Appl. Phys. Lett. **68**, 514 (1996); M. Asif Khan, Q. Chen, C. J. Sun, M. Shur, and B. Gelmont, *ibid.* **67**, 1429 (1995).

⁹S. N. Mohammad, Z.-F. Fan, A. Salvador, O. Aktas, A. E. Botchkarev, W. Kim, and Hadis Morkoc, Appl. Phys. Lett. **69**, 1420 (1996).

¹⁰C. Johnson, J. Y. Lin, H. X. Jiang, M. Asif Khan, and C. J. Sun, Appl. Phys. Lett. **68**, 1808 (1996).

¹¹S. Nakamura, T. Mukai, and M. Senoh, J. Appl. Phys. **71**, 5543 (1992).

¹²D. E. Lacklison, J. J. Harris, C. T. Foxon, J. Hewett, D. Hilton, and C. Roberts, Semicond. Sci. Technol. **3**, 633 (1988).

¹³T. N. Theis and S. L. Wright, Appl. Phys. Lett. **48**, 1374 (1986).

¹⁴A. Dissanayake, M. Elahi, H. X. Jiang, and J. Y. Lin, Phys. Rev. B **45**, 13 996 (1992); J. Y. Lin, A. Dissanayake, G. Brown, and H. X. Jiang, *ibid.* **42**, 5855 (1990).

¹⁵H. J. Quieser and D. E. Theodorou, Phys. Rev. B **33**, 4027 (1986).

¹⁶H. L. Störmer, A. C. Gossard, G. Wiegmann, and K. Baldwin, Appl. Phys. Lett. **39**, 912 (1981).

¹⁷B. Saffian, W. Kraak, B. Oelze, H. Kunzel, and J. Bottcher, Phys. Status Solidi B **196**, 323 (1996).

¹⁸P. M. Mooney, J. Appl. Phys. **67**, R1 (1990).

Microstructure, vibrational and electronic properties of GaN grown by molecular beam epitaxy on Al₂O₃(0001) and 6H-SiC(0001)

D. Freundt, D. Holz, and H. Lüth

Institut für Schicht und Ionentechnik (ISI), Forschungszentrum Jülich, D-52425 Jülich, Germany

M. Romani and A. Rizzi

INFN-Dipartimento di Fisica, Università di Modena, via Campi 213/A, I-41100 Modena, Italy

D. Gerthsen

Laboratorium für Elektronenmikroskopie, Kaiserstrasse 12, D-76128 Karlsruhe, Germany

(Received 12 January 1997; accepted 9 May 1997)

GaN layers have been grown by molecular beam epitaxy with a rf plasma source on Al₂O₃(0001) and 6H-SiC(0001). The conductive *n*-SiC substrates were employed for the *in situ* characterization of the grown GaN layers by electron spectroscopies (HREELS and XPS). Transmission electron microscopy (TEM) in the conventional and high-resolution mode provides information regarding the structural properties. Plan-view TEM yields a threading defect density of $7 \times 10^9 \text{ cm}^{-2}$ in GaN/6H-SiC and $2 \times 10^{10} \text{ cm}^{-2}$ in GaN/Al₂O₃. Micro-Raman spectroscopy, by analysis of the coupled A₁(LO)-phonon-plasmon mode, provides the free carrier concentrations to be $n \sim 1 - 2 \times 10^{17} \text{ cm}^{-3}$ for GaN layers grown on both substrates. The *in situ* HREEL spectroscopy gives insight into the electronic and vibrational properties of the GaN surface. Deep levels at $\sim 900 \text{ meV}$ above the valence band maximum induce a surface absorption structure which is reduced after heating at 600°C and which might be assigned, according to the growth process and to the literature, to the presence of Ga vacancies in the as-grown layers, in particular in the region close to the surface. Heating the GaN/6H-SiC heterostructures at 900°C causes complete decomposition of the GaN layer. © 1997 American Vacuum Society. [S0734-211X(97)12104-7]

I. INTRODUCTION

The III-V nitrides are presently the most promising material system for optoelectronic applications in the green to UV spectral range¹ and for superior high-voltage unipolar power devices operating at high temperatures.² The direct band gap of the wurtzitic polytypes varies from 1.9 eV for InN, 3.4 eV for GaN, to 6.2 eV for AlN. High-brightness LEDs on the basis of (In, Al)GaN alloyed heterostructures have been fabricated on sapphire¹ and SiC.³ The crystal growth of these materials is still a major problem in the realization of nitride based devices. The lack of commercially available bulk material demands heteroepitaxy on lattice and thermally mismatched substrates. The nitrides are mostly grown by slightly different modifications of either metal-organic vapor phase epitaxy (MOVPE), metalorganic chemical vapor deposition (MOCVD) or molecular beam epitaxy (MBE). Since the introduction of low radiation damage nitrogen-plasma sources for the growth of nitrides, in particular radio frequency plasma sources,⁴ there are new possibilities for MBE grown material. Furthermore, MBE processes based on thermal cracking of ammonia (NH₃) have also been shown to be successful.⁵ The technological relevance of the group III nitrides for the semiconductor industry requires further improvement of their structural and optical properties and a detailed understanding of the correlation between these properties and the growth process. Comparatively little effort has been devoted to the surface and interface structure of these materials, though both aspects are of central importance for the development of a nitride-based technology. MBE offers

the possibility of *in situ* analysis and the GaN surface electronic, vibrational, and structural properties were investigated in this work with high resolution electron energy loss spectroscopy (HREELS), x-ray photoemission spectroscopy (XPS), and low energy electron diffraction (LEED). *Ex situ* characterization of the MBE grown layers on sapphire and 6H-SiC, was performed by micro-Raman spectroscopy and transmission electron microscopy (TEM) in the conventional and high resolution mode.

II. EXPERIMENT

The samples are grown in a MBE system that is equipped with standard effusion cells for the group III elements (Ga 7N) and a rf plasma source (SVTA Model-rf4.5)⁴ to provide a flux of nitrogen radicals. High purity nitrogen (6N) is introduced through a mass flow controller into the plasma source. The samples for the present investigation are all grown with a flow rate of 2 sccm and a rf power of 350 W. The base pressure of the MBE system is $5 \times 10^{-11} \text{ mbar}$ and rises up to $3.5 \times 10^{-5} \text{ mbar}$ during the epitaxy. The MBE is connected with two analysis chambers. The first one is equipped with a three-grid LEED system, a CMA for AES and a spectrometer for the energy loss experiments (HREELS) with two stage monochromator and analyzer (EELS90). The HREEL spectra are taken with primary electron energies of 7 and 20 eV and a resolution of 3.5 meV for the measurements in the infrared energy region and with 9.5 meV resolution for the measurements in the electronic transition energy region. The base pressure in this chamber is

1×10^{-11} mbar. A second analysis chamber is devoted to photoemission spectroscopy with a monochromatized x-ray source and a half-sphere Leybold analyzer. The overall resolution of the photoemission spectra ranges from 0.4 to 1.6 eV, depending on the pass energy of the analyzer.

Al_2O_3 and 6H-SiC wafers with (0001) orientation are used as substrate materials. Prior to growth the sapphire wafers are degreased in organic solvents. Heating in UHV for 15 min at 750 °C results in a (1×1) surface without any carbon contamination, as checked with AES. The SiC wafers are degreased in organic solvents, then etched with $\text{HCl}:\text{H}_2\text{SO}_4$ (1:1), $\text{NH}_3:\text{H}_2\text{O}_2:\text{H}_2\text{O}$ (5:3:3), $\text{HCl}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$ (5:3:3) and eventually dipped in buffered ammonium fluoride etchant (AF91). A thin oxide layer is grown at a temperature of 1200 °C with an oxygen flow rate of 10 sccm. Before insertion into UHV the SiC wafers are etched with 10% HF. Heating in UHV at 750 °C for 15 min results in a sharp (1×1) surface with only a very small amount of oxygen, that is slightly above the detection limit of AES.

The growth of GaN on Al_2O_3 starts with the nitridation of the surface, which is exposed to the flux of nitrogen radicals at a substrate temperature of 550 °C for 30 min. At the same temperature a 20 nm thick GaN buffer layer is grown with a growth rate of 40 nm/h.⁶ The temperature is then increased up to 760 °C for the growth of the epitaxial GaN film with a growth rate of 300 nm/h. The growth of GaN on SiC follows the same procedure with exception of the initial nitridation step.

First *ex situ* characterization of the grown GaN layers is done by high-resolution x-ray diffraction. The $\text{Cu } K\alpha_1$ line ($\lambda = 0.1541$ nm) is used as the x-ray probe. Photoluminescence spectra are taken with a frequency-doubled Ar laser at a wavelength of 245 nm and a power of 2.5 mW. Raman scattering measurements are performed in backscattering geometry using a micro-Raman setup. The excitation source is the 457.9 nm line of an Ar laser with a power of 20 mW. The scattered light is analyzed with a triple monochromator (Dilor XY) and is detected by an optical multichannel analyzer that consists of a microchannel plate as photomultiplier and a linear silicon photodiode array. The Raman spectra are all taken at room temperature.

The nanostructural properties of the GaN layers were investigated by conventional and high-resolution transmission electron microscopy (TEM) and (HRTEM). The images were taken with a Philips CM200 FEG/ST electron microscope with a Scherzer resolution of 0.24 nm and the atomic resolution electron microscope (Max-Planck-Institut für Metallforschung, Stuttgart/Germany) with a Scherzer resolution of 0.12 nm with the analytical objective polepiece.

A general assessment of the quality of the MBE GaN epitaxial layers whose results are presented and discussed in the next sections is given here, based on XRD and PL results. The rocking curve for the best GaN layers grown so far on Al_2O_3 (0001) show a FWHM of 7 arcmin for the (0002) reflex. The use of a triple-axis analyzer between the sample and the detector gives a FWHM for the ω scan that reflects the width of the rocking curve, a hint of regions in the

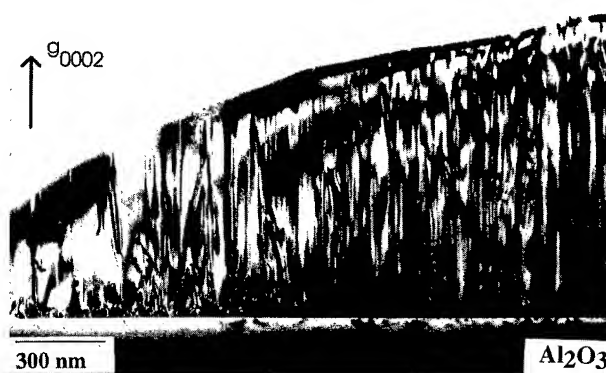


FIG. 1. TEM conventional images of GaN on Al_2O_3 (0001) using imaging vector $\mathbf{g} = (0002)$.

sample which are tilted with respect to each other (mosaic spread). PL measurements at 2 K show only a peak at 3.467 eV with a FWHM of 4.5 meV which is attributed to the donor-bound exciton (D^0, X). At these low temperatures there are no traces of the "yellow luminescence."

III. STRUCTURAL PROPERTIES BY TEM

Figure 1 shows a conventional TEM image of GaN on Al_2O_3 (0001) using a $\mathbf{g} = (0002)$ imaging vector. Two different types of threading defects can be distinguished. The first type runs straight from the interface to the surface which dominates the defect structure. It exhibits a bandlike contrast with a width of approximately 6–10 nm. The density does not significantly decrease with increasing distance from the interface. The majority of the bands disappear if an imaging vector $\mathbf{g} = (1\bar{1}00)$ is used. Using convergent beam electron diffraction, Daudin *et al.*⁷ identified similar looking defects as inversion domain boundaries. However, the authors do not present a conventional contrast analysis. The density of threading dislocations which remain under $\mathbf{g} = (1\bar{1}00)$ imaging conditions is distinctly reduced with increasing layer thickness. The same defect types are observed in GaN on 6H-SiC(0001). Plan-view TEM shows that the threading defect density in the upper part of the GaN layer is $7 \times 10^9 \text{ cm}^{-2}$ in GaN/SiC and about $2 \times 10^{10} \text{ cm}^{-2}$ in GaN/ Al_2O_3 . Despite the significantly smaller misfit of GaN/6H-SiC(0001) the defect density is in the same order of magnitude.

Figure 2 shows HRTEM cross-section images along the $[1\bar{1}00]_{\text{GaN}}$ projection on Al_2O_3 [Fig. 2(a)] and on SiC [Fig. 2(b)]. This GaN projection is rarely imaged because the distance of the $(1\bar{1}20)$ -lattice fringes of 0.1595 nm perpendicular to the interface is difficult to resolve in many transmission electron microscopes. The images reveal significant differences between the buffer layer structures on SiC and Al_2O_3 . The GaN/ Al_2O_3 (0001) interface is flat on an atomic level over some 10 nm. Interface steps [one example being marked by an arrow in Fig. 2(a)] with a height of $1/6c_{\text{Al}_2\text{O}_3}$ corresponding to 0.2166 nm separate flat sections. The (0002)-lattice fringes in the GaN are bent above the step

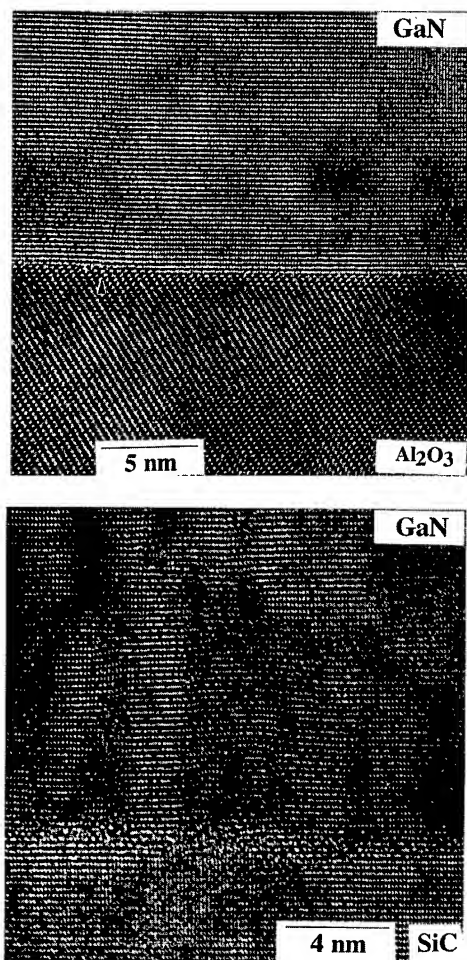


FIG. 2. Cross-sectional HRTEM images along the $[1\bar{1}00]_{\text{GaN}}$ projection for GaN on $\text{Al}_2\text{O}_3(0001)$ (a, top) and on $6\text{H-SiC}(0001)$ (b, bottom). The arrow in (a) marks a substrate step of the height of $1/6c_{\text{Al}_2\text{O}_3}$.

which is indicative of high stress. The bending occurs because a mismatch exists between the $(0002)_{\text{GaN}}$ planes with a distance of 0.2594 nm and the smaller step height of the Al_2O_3 . The interface of the GaN on 6H-SiC is distinctly less perfect. The observed disturbances can partly be attributed to the surface preparation prior to the growth and partly to the radiation damage due to the 1.25 MeV electrons. Small domains with extensions of some nanometers can be recognized which are separated by planar defects induced at the interface. A number of different models for the planar faults were already presented. The faults can be subdivided into inversion domain boundaries with or without a translation of $c_{\text{GaN}}/2$.⁸ The second class involves boundaries without a polarity inversion with different translation vectors, e.g., stacking mismatch boundaries in GaN/ $6\text{H-SiC}(0001)$.⁹ The shift of the (0002) fringes in the GaN in Fig. 2(b) and the contrast extinction for $g=(1\bar{1}00)$ indicates that a translation along the c -axis occurs across the boundaries. The polarity cannot be determined from a single HRTEM image.

HRTEM was also carried out along the $[11\bar{2}0]_{\text{GaN}}$ projection which encloses an angle of 30 degrees with the $[1\bar{1}00]_{\text{GaN}}$ projection. The domain structure is more difficult

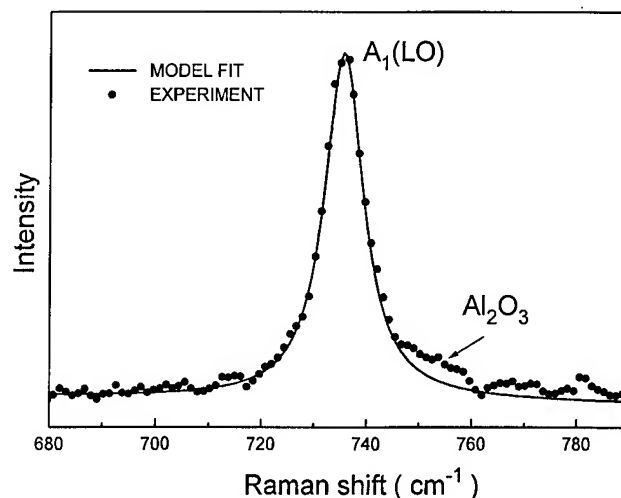


FIG. 3. Micro-Raman spectrum (dots) of a GaN layer ($\sim 1.5 \mu\text{m}$ thick) on sapphire in the region of the GaN $A_1(\text{LO})$ mode. The spectrum has been measured in the backscattering configuration $z(xx)z$ for which the $A_1(\text{LO})$ mode is permitted. The measured structure is a longitudinal-phonon-plasmon (LPP) coupled mode, due to the unintentional n doping of the as-grown GaN layers. The continuous line is a model calculation of the coupled mode. The parameters of the model are given.

to detect because a high density of stacking faults parallel to the substrate is also visible. The density of the stacking faults parallel to the substrate is higher on the SiC than on the Al_2O_3 substrate. At a distance of more than 100 nm from the interface, the density of planar faults strongly decreases. The high defect density in the GaN buffer on the SiC is likely to be correlated with the comparatively high defect density in the upper part of the GaN layer.

Small inclusions of zinc-blende GaN with extensions of only a few nanometers are observed on Al_2O_3 if the $[11\bar{2}0]_{\text{GaN}}$ projection is used. Cubic GaN was previously observed in MOCVD-grown low-temperature nucleation layers on $\text{Al}_2\text{O}_3(0001)$ by Wu *et al.*¹⁰ The presence of zinc-blende material indicates that the material is not in the state of thermodynamic equilibrium. However, only speculations are possible about the origin of the cubic GaN. The strain induced pressure that results from the difference of the thermal expansion coefficients between GaN and Al_2O_3 could lead to the local phase transformations from the wurtzite into the zinc-blende structure. The presence of strain is confirmed by the results of the Raman measurements. Another possibility could be deviations from the proper stoichiometry which also might induce high local pressures in the layer.

IV. MICRO-RAMAN SPECTROSCOPY

The wurtzite structure belongs to the space group C_{6v}^4 ($C6_3mc$) with two formula units per primitive cell. Group theory predicts the following lattice phonons: A_1 and E_1 polar modes, polarized parallel to the optical axis and in the basal plane, respectively, both infrared and Raman active; two E_2 and two B_1 branches, the first Raman active, the latter silent.¹¹ The anisotropy of GaN in the short-range atomic forces is responsible for the A_1-E_1 splitting, while

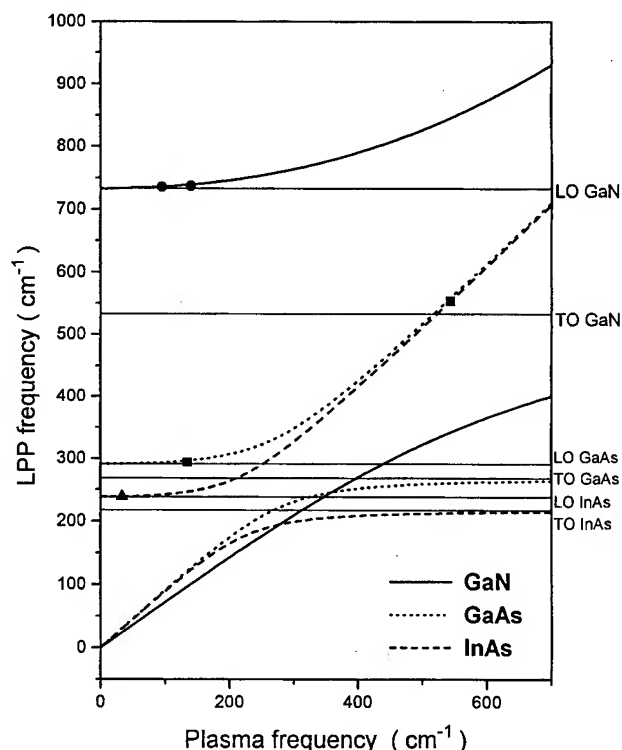


FIG. 4. Calculated LPP coupled modes in GaN, GaAs, and InAs as a function of the plasma frequency. The experimental points are determined by the fit of the LPP Raman spectrum. The horizontal lines corresponds to the LO and TO phonon energies in the different materials.

the long-range Coulomb field arising from the partial ionic character of the crystal bonding is the origin of the longitudinal-transverse (LO-TO) splitting of the polar modes.

The grown GaN layers were characterized by micro-Raman spectroscopy to probe their crystallinity and to determine the unintentional free carrier concentration. Raman selection rules for the wurtzite structure give vanishing intensity of the $A_1(\text{TO})$ mode in the $z(xx)z$ backscattering configuration. The crystal disorder related to the mosaic structure of the grown GaN layers is reflected in a measurable intensity of the forbidden mode. A clear correlation is seen between the intensity of the forbidden mode and the width of the rocking curve, namely decreasing intensity with decreasing XRC-FWHM.

GaN layers grown epitaxially on different substrates normally show an unintentional n doping, generally assigned to the presence of N vacancies. In a polar semiconductor the free carrier plasma interacts with the LO phonons via their macroscopic electric fields giving rise to the phenomenon of "level anticrossing." Two coupled modes appear, which have mixed plasmon-phonon character (LPP modes). From a model fit of the higher-frequency coupled mode, the electron concentration can be optically determined without the need of contacting the sample and with high lateral resolution.¹² The line shape of the coupled mode according to Imer et al.¹² is essentially given by the deformation potential and by the macroscopic longitudinal electric field being predomi-

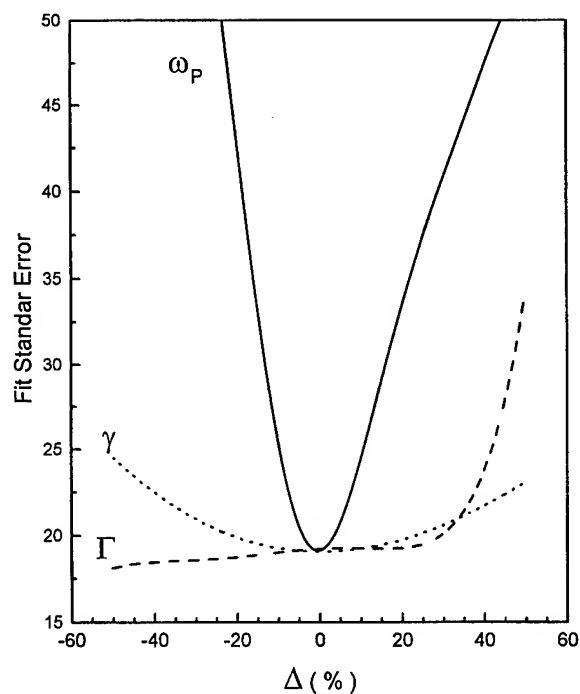


FIG. 5. Fit standard error as a function of the percentage variation from the best value of the parameters used to fit the Raman spectrum. ω_p is the plasma frequency, γ and Γ are the plasmon and $A_1(\text{LO})$ phonon damping constants, respectively.

nant in GaN, as compared with scattering from density fluctuations of the free carriers.¹³ The material parameters entering the model are the frequencies $\omega_T = 533 \text{ cm}^{-1}$ and $\omega_L = 733 \text{ cm}^{-1}$ of the TO and LO phonons, the plasmon (γ) and phonon (Γ) damping constants, the plasma frequency ω_p (and therefore the free carrier concentration, the high frequency dielectric constant, and the effective mass) and the Faust-Henry coefficient $C = 0.43$. The procedure used for the present fit resembles that in the work of Harima et al.¹⁴ on SiC. Taking ω_p , γ , and Γ as fitting parameters, the measured $A_1(\text{LO})$ LPP mode was fitted to the calculated line shape. A representative spectrum is shown in Fig. 3. The calculation well reproduces the experimental line shape with fitting parameters given by $\omega_p = 95.1 \text{ cm}^{-1}$, $\gamma = 306.2 \text{ cm}^{-1}$, and $\Gamma = 6.5 \text{ cm}^{-1}$. On the high energy side the disagreement is due to some signal arising from the Al_2O_3 substrate (E_g at 750 cm^{-1}), as indicated by the arrow. This method for the determination of the plasma frequency and therefore of the carrier concentration in the grown layers was checked by comparison with Hall results on n -GaAs and n -InAs MBE layers. In Fig. 4 the calculated LPP branches are plotted against the plasma frequency for GaN, GaAs, and InAs. The two coupled modes are obtained by solving the equation $\epsilon(\omega) = 0$, where $\epsilon(\omega)$ is the total dielectric function of the semiconductor in the presence of the free carriers (Lorentz oscillator and Drude model).¹⁵ The points are obtained from the fit of the Raman spectra in the region of the $A_1(\text{LO})$ (higher coupled mode) for GaN and of the LO mode (higher coupled branch) for the other zinc-blende semiconductors. The model fit provides the plasma frequency corre-

TABLE I. Electron concentrations for different samples of GaAs and InAs determined by Hall effect and by fitting of the Raman LPP⁺ coupled mode. The GaN free carrier concentrations have been determined only by the optical method.

	Hall n (cm ⁻³)	Raman n (cm ⁻³)
GaAs (No. 1)	2.4×10^{18}	2.4×10^{18}
GaAs (No. 2)	3.15×10^{17}	1.5×10^{17}
InAs	2.2×10^{17}	1.0×10^{17}
GaN/Al ₂ O ₃ (No. 1)		2.2×10^{17}
GaN/Al ₂ O ₃ (No. 2)		1.2×10^{17}

sponding to the measured LPP⁺ frequency; the fact that the points lie on the calculated branches is inherent in the followed procedure. The point here is to show that with the high energy resolution of the Raman measurements (1 cm⁻¹), even if the range of interest for the GaN is on the flatter part of the curve, the error in ω_p is estimated to be within 5% (Fig. 5). In this plot the standard error of the fit procedure is represented as a function of the percentage variation from the best value of the free parameters, ω_p , γ , and Γ . The fit of the Raman line shape shows no significant dependence on variations of the damping constants and therefore the free carrier mobility for these samples cannot

be determined in a plausible way by this method. For the sake of comparison, Table I shows the electron concentrations determined by Hall effect and by the described optical method for GaAs and InAs epitaxial layers. The derived electron concentration for different GaN layers grown on sapphire is $n \sim 1-2 \times 10^{17}$ cm⁻³. Compared with Hall measurements this optical technique for the determination of the free carrier concentration separates the contribution of different conductive materials in a heterostructure and therefore can be conveniently applied, e.g., in the presence of other conductive layers or of a conductive substrate.

V. THERMAL TREATMENT OF GaN: HREELS AND XPS

The vibrational and electronic surface properties of GaN/6H-SiC heterostructures have been investigated *in situ* by HREELS and XPS. In particular, the surface of the *as-grown* GaN epilayers (~ 300 nm of thickness) and after successive heating steps in UHV at 600, 700, 800, and 900 °C (10 min each) was characterized. Figure 6(a) shows the HREEL spectra measured in the energy loss range of the electronic excitations across the GaN band gap. The onsets of two types of transitions are clearly recognized. Both structures have been fitted with a $(h\nu - E)^{1/2}$ type of function reflecting the probability of absorption for allowed band-to-band direct transitions and for electron transitions from an impurity level into

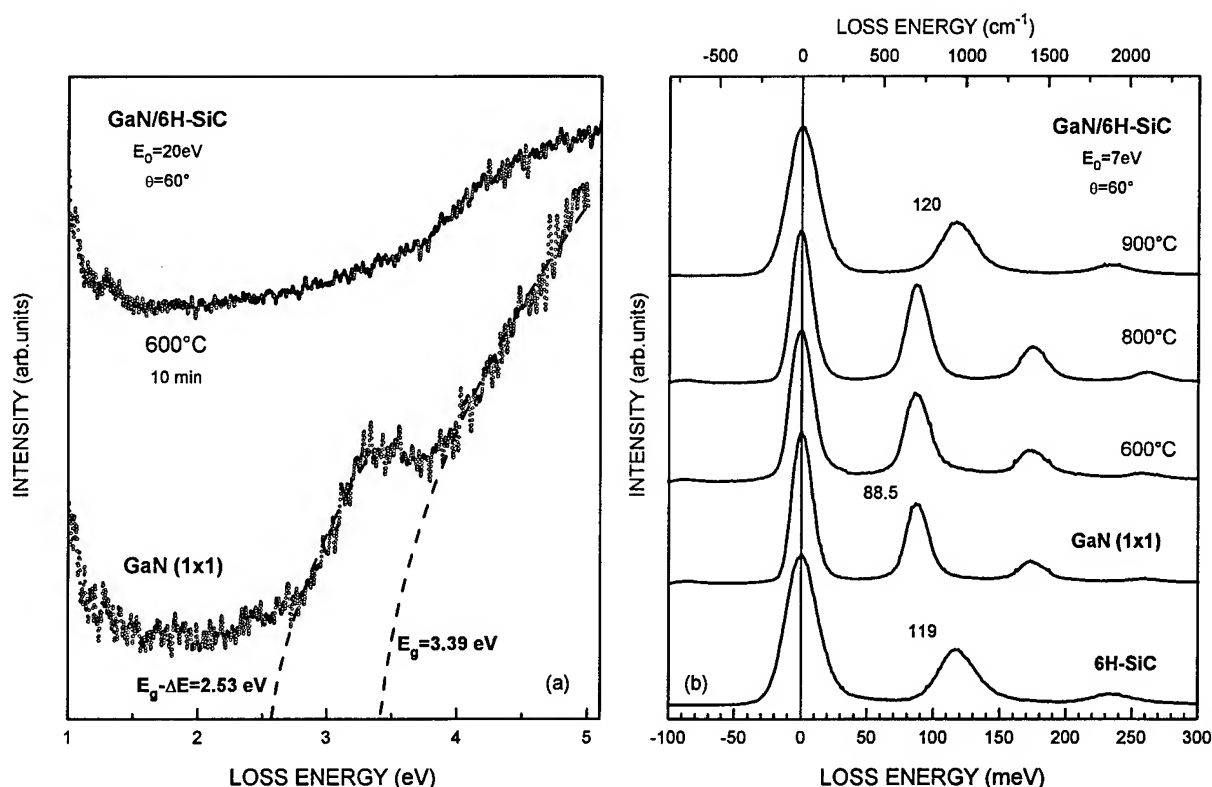


FIG. 6. HREEL spectra measured at RT at the surface of the *as-grown* GaN/6H-SiC(0001) heterostructures and after successive annealing steps. (a) HREEL spectra taken in the energy range of the electronic excitations across the GaN band gap. The spectra are normalized to the quasi-elastic peak (not shown) and represented with a magnification factor of 10^4 on a full-scale elastic-peak height. (b) HREEL spectra in the IR energy range together with the quasi-elastic peak at zero loss energy. The spectrum at the bottom is that of the 6H-SiC substrate surface.

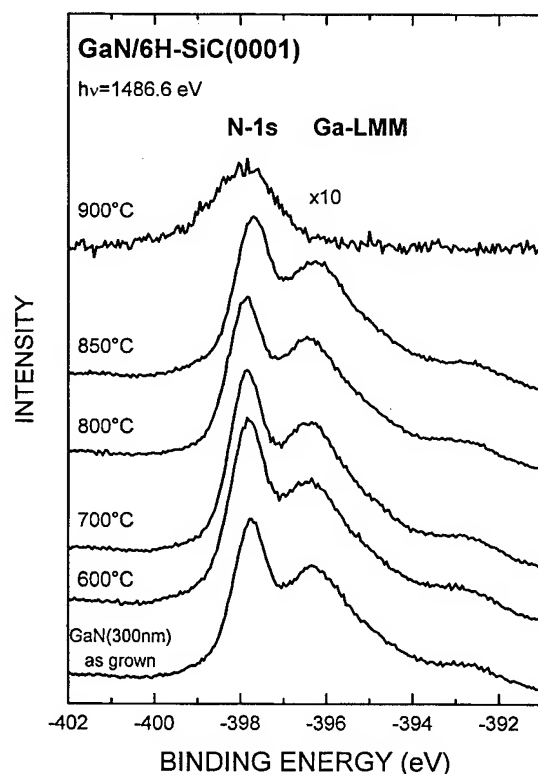


FIG. 7. XPS spectra of GaN grown on 6H-SiC(0001) in the binding energy region of N 1s core level and Ga LMM Auger excitations. The thickness of the epilayer is 300 nm. The spectra are measured on the as-grown sample (bottom) and after successive annealing steps in UHV at the indicated temperature. The last spectrum at the top is plotted with a magnification factor of 10.

the conduction band.¹⁶ The higher loss structure corresponds to excitations of electrons across the energy gap of the GaN, $E_g = 3.39$ eV, the lower one to excitations from a deep energy level in the gap into the conduction band with $\Delta E = 860$ meV above the valence band maximum. After heating at 600 °C the lower absorption band is strongly reduced. In Fig. 6(b) the HREEL spectra are shown in the IR energy loss range. The observed strong features are due to the surface Fuchs-Kliwer phonons (FK phonons) and their multiple excitations. The bottom spectrum is that of the 6H-SiC clean surface. The commercial SiC wafers are *n* doped with a nominal electron concentration of $1 \times 10^{18} \text{ cm}^{-3}$ and this is reflected in the broad FWHM of the elastic peak (FWHM = 32.8 meV) to be compared with the experimental energy resolution of 3.5 meV. A further reason for the broadening of the elastic peak is also a certain roughness of the surface on a scale of 8 nm, which has been revealed by atomic force microscopy. The FK phonons of the *as-grown* GaN surface are found at 88.5 meV. The FWHM of the quasi-elastic peak is 17 meV, due to the presence of free carriers as well as a certain surface roughness. Heating of the sample up to 800 °C does not induce any change in the IR surface properties; after the heating step at 900 °C the HREEL spectrum resembles that of the substrate, with the FK phonons of SiC at 119 meV and with the broader quasi-elastic peak.

The GaN MBE growth process runs in a nitrogen rich

atmosphere and at the end of the growth the Ga source is first closed, keeping the N source running to avoid a nitrogen depletion in the surface near region. The GaN surface always shows a (1×1) reconstruction with a light diffused intensity background, as seen in LEED. The electronic structure and the formation energies of native defects in GaN have been calculated and the dominant intrinsic point defects are found to be vacancies.¹⁷ In *n*-type GaN the Ga vacancy gives rise to a deep acceptor level (a triple acceptor in a three charge state). The relevant transition level is calculated to lie at ~ 1.1 eV above the top of the valence band; therefore Ga vacancies and their complexes with O are inferred to be responsible for the yellow luminescence observed in *n*-GaN.¹⁸ The energy position of these deep acceptor levels is also consistent with the lower absorption band observed in HREELS. The N-rich growth conditions, in particular at the end of the growth suggest that Ga vacancies might increasingly be formed in a region close to the surface. The behavior of the HREEL spectra in the vibrational region is very clear: it shows the thermal decomposition of the GaN layer in UHV.

This result is further confirmed by the XPS spectra of Fig. 7. The structure observed at approximately -398 eV binding energy is due to the N 1s photoelectrons; close by at lower binding energies the Ga LMM Auger excitations are measured. The escape depth for both lines therefore amounts to the same value, $\lambda \sim 1.8$ nm. No significant variation of the relative intensity is observed after each annealing step, up to 850 °C. Annealing at 900 °C causes the complete desorption of GaN (also revealed by the disappearance of all the other GaN related XPS core level structures); the small N 1s signal left is assigned to a reacted component with the SiC substrate. The shift of 0.1 eV towards lower binding energy after 850 °C is observed for all XPS structures and is therefore assigned to a change in the band bending at the surface.

The XPS results clearly show that Ga and N are desorbed together from the GaN surface; the thickness of the epilayer might gradually decrease after each annealing step but this behavior cannot be followed by XPS and HREELS, which are surface sensitive techniques. These results are not in contradiction with the conclusion of a congruent thermal decomposition in vacuum due to the large kinetic barrier of GaN decomposition as reported in the literature.¹⁹ Nevertheless a mass spectroscopic study is needed to support this conclusion.

VI. CONCLUSIONS

Epitaxial GaN layers grown on $\text{Al}_2\text{O}_3(0001)$ and 6H-SiC(0001) by plasma assisted MBE have been characterized by several techniques. The TEM analysis shows a similar defect structure for the layers grown on both substrates. Inversion domain boundaries with a bandlike contrast, threading dislocations, and planar defects in the region close to the interface are found. A higher density of planar defects was observed in GaN on 6H-SiC. This can be considered as a

possible origin of the threading defect density in the upper part of the GaN layer which is reduced only by a factor 3 compared to GaN/Al₂O₃. Micro-Raman spectroscopy has been employed to determine the unintentional doping concentration. The *in situ* HREEL spectroscopy gives insight into the electronic and vibrational properties of the GaN surface. Deep levels at ~900 meV above the valence band maximum induce a surface absorption structure which might be attributed to the presence of Ga vacancies in the as-grown layers, in particular in the region close to the surface. Annealing the GaN/6H-SiC heterostructures at 900 °C causes decomposition of the GaN layer.

ACKNOWLEDGMENTS

Technical support by F. Ringelmann is gratefully acknowledged; we thank D. Guggi and G. Crecelius for the XRD and PL characterization of the layers, respectively. The authors would like to thank W. Sigle, H. Höschen, and F. Phillipp at the Max-Planck-Institut für Metallforschung Stuttgart/Germany for the opportunity to use the atomic resolution electron microscope and the support during the experiments.

¹S. Nakamura, T. Mukai, and M. Senoh, Appl. Phys. Lett. **64**, 1687 (1994).

²T. P. Chow and R. Tyagi, IEEE Trans. Electron Devices **41**, 1481 (1994).

³Data sheet from CREE Research Inc., 1810 Meridian Parkway, Durham, NC 27713 (1995).

⁴W. E. Hoke, P. J. Lemonias, and D. G. Weir, J. Cryst. Growth **111**, 1025 (1991); R. P. Vaudo, Z. Yu, J. W. Cook, Jr., and J. F. Schetzina, Opt. Lett. **18**, 1843 (1993); J. M. Van Hove, G. Carpenter, E. Nelson, A. Wowchak, and P. P. Chow, J. Cryst. Growth **164**, 154 (1996).

⁵W. Kim, Ö. Aktas, A. E. Botchkarev, A. Salvador, S. N. Mohammad, and H. Morkoç, J. Appl. Phys. **79**, 7657 (1996).

⁶S. Nakamura, Jpn. J. Appl. Phys. **30**, L1705 (1991).

⁷B. Daudin, J. L. Rouviere, and M. Arlery, Appl. Phys. Lett. **69**, 2480 (1996).

⁸P. Vermaut, P. Ruterana, G. Nouet, and H. Morkoç, Philos. Mag. A **75**, 239 (1997); P. Vermaut, P. Ruterana, G. Nouet, A. Salvador, and H. Morkoç, Inst. Phys. Conf. Ser. (in press).

⁹B. N. Sverdlov, G. A. Martin, H. Morkoç, and D. J. Smith, Appl. Phys. Lett. **67**, 2063 (1995).

¹⁰X. H. Wu, D. Kapolnek, E. J. Tarsa, B. Heying, S. Keller, B. P. Keller, U. K. Mishra, S. P. DenBaars, and J. S. Speck, Appl. Phys. Lett. **68**, 1371 (1996).

¹¹C. A. Arguello, D. L. Rousseau, and S. P. S. Porto, Phys. Rev. **181**, 1351 (1969).

¹²G. Irmer, V. V. Toporov, B. H. Batramov, and J. Monecke, Phys. Status Solidi B **119**, 595 (1983).

¹³T. Kozawa, T. Kachi, H. Kano, Y. Taga, M. Hashimoto, N. Koide, and K. Manabe, J. Appl. Phys. **75**, 1098 (1994).

¹⁴H. Harima, S. Nakashima, and T. Uemura, J. Appl. Phys. **78**, 1996 (1995).

¹⁵P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 1996), pp. 282–300; H. Ibach and H. Lüth, *Solid-State Physics* (Springer, Berlin, 1991), pp. 247–262.

¹⁶R. A. Smith, *Semiconductors*, 2nd ed. (Cambridge University Press, London, 1978), pp. 309–337.

¹⁷J. Neugebauer and C. G. Van de Walle, Phys. Rev. B **50**, 8067 (1994).

¹⁸J. Neugebauer and C. G. Van de Walle, Appl. Phys. Lett. **69**, 503 (1996).

¹⁹Z. A. Munir and A. W. Searcy, J. Chem. Phys. **42**, 4223 (1965).

Raman monitoring of molecular beam epitaxial growth of GaN on GaAs (100) and Si (111)

D. Drews,^{a)} A. Schneider, and D. R. T. Zahn

Professur für Halbleiterphysik der TU Chemnitz-Zwickau, D-09107 Chemnitz, Germany

(Received 12 January 1997; accepted 23 March 1997)

Epitaxial layers of GaN were grown by molecular beam epitaxy on GaAs (100) and Si (111) substrates. Nitrogen was provided from a rf-plasma source while elemental Ga was evaporated from a Knudsen cell. The growth process was performed in an ultrahigh vacuum chamber that is optically aligned with a multichannel Raman spectrometer. This setup allows Raman spectra to be taken *online*, i.e., during the growth process. Utilizing resonant Raman scattering conditions, spectra with a sufficient signal-to-noise ratio were taken even at the growth temperature of 600 °C. For both substrates the evolution of compressive strain at the interface was monitored from the frequency shift of the substrate phonons in the initial phase of the growth process. © 1997 American Vacuum Society. [S0734-211X(97)07104-7]

I. INTRODUCTION

The success in the fabrication of GaN-based optoelectronic devices operating in the blue and ultraviolet spectral region, e.g., light-emitting diodes and injection lasers,^{1,2} has launched an enormous research interest in the basic material properties as well as the understanding of the growth mechanisms for heterostructures. Detailed knowledge in these fields is expected to help improve the quality of the material and the performance of devices. Epitaxial growth of GaN layers has been achieved by metalorganic vapor phase epitaxy (MOVPE)³ and molecular beam epitaxy (MBE)⁴ on various substrates, including sapphire, GaAs, and Si, but severe problems result from the high growth temperatures around 1000 °C in MOVPE and up to 700 °C in MBE as well as the high lattice mismatch (GaAs: $\Delta a/a = 0.2$, Si: $\Delta a/a = 0.17$ for cubic GaN at RT). Growth techniques have to be further improved in order to cope with these obstacles, and consequently there is a need for sophisticated characterization tools that can be applied during growth in order to obtain direct information on the growth process. While reflection high energy electron diffraction is commonly used in MBE, several optical techniques have been developed in recent years, the application of which is not restricted to ultrahigh vacuum (UHV) conditions. Examples are reflectance anisotropy spectroscopy and spectroscopic ellipsometry.⁵ Also Raman scattering, which has a long tradition as a characterization method for thin films and heterostructures, has been successfully applied for monitoring the MBE growth of II–VI compounds such as ZnSe⁶ or ternary compounds like ZnS_xSe_{1-x}.⁷ Information deduced from these experiments includes layer composition, layer thickness, strain evolution, crystal quality, and interface reactivity.⁸ Although the scattering volume of thin films is very small, sufficient sensitivity for the detection of ultrathin layers can be achieved by taking advantage from resonant Raman scattering (RRS) that leads to a strong enhancement of the scattering cross section when the excitation energy is chosen in the vicinity of the

fundamental band gap or another critical point in the band structure of the material under investigation.⁹ For growth monitoring this condition has to be fulfilled at the growth temperature, so the temperature dependence of the band gap has to be taken into account.

In the present article we report on Raman monitoring of the MBE growth of GaN on GaAs (100) and Si (111) substrates.

II. EXPERIMENT

MBE growth was performed in a specially designed UHV chamber (base pressure $< 10^{-10}$ mbar) that allows light scattered from the sample to be collected through the central viewport and focused on the entrance slit of a Raman spectrometer (Dilor XY) with multichannel detection. Details of this setup can be found in Ref. 10. Elemental Ga (6N purity) was evaporated from a Knudsen cell at a temperature of 870 °C as measured by a thermocouple. This leads to a Ga flux of about 5×10^{12} atoms/cm² s as judged from the frequency decrease rate of a quartz microbalance. Atomic nitrogen was provided by a rf-plasma source (SVT RF 4.5) operated at a power of 300 W and a nitrogen (6N purity) flux of 2 sccm leading to a total pressure around 2×10^{-4} mbar in the chamber during growth. The growth was performed at a substrate temperature of 600 °C, which was determined from the frequency shift of the substrate phonons. The GaAs (100) substrate surface was prepared by thermal desorption of an arsenic cap deposited on a micron thick GaAs homoepitaxial layer grown by MBE. This resulted in a clean As rich surface with $c(4 \times 4)$ reconstruction as observed by low-energy electron diffraction. The Si (111) substrate surface was prepared by chemically etching and subsequent annealing at 825 °C in UHV. The 406.7 nm (3.05 eV) emission line of a Kr⁺ ion laser was used for excitation, and the monochromator slits were set for a resolution of 5.5 cm⁻¹. The laser beam ($P < 100$ mW) was focused to a spot of about 100 μ m in diameter on the sample surface. All Raman spectra were recorded in a backscattering geometry, the polarization of the incident light was parallel to the [011] direction of the sub-

^{a)}Corresponding author; Electronic mail: drews@physik.tu-chemnitz.de

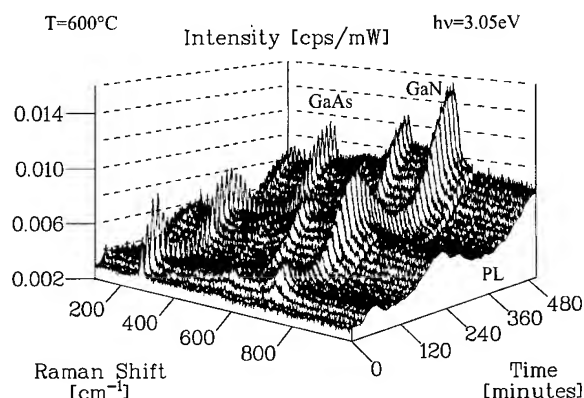


FIG. 1. Evolution of Raman spectra during growth of GaN on GaAs (100) at 600 °C ($h\nu = 3.05$ eV).

strate. For the sake of higher sensitivity the polarization of the scattered light was not analyzed, but the detection of light with parallel polarization is strongly favored by the monochromator response.

III. RESULTS AND DISCUSSION

Figure 1 shows a three-dimensional plot of the evolution of Raman spectra during the growth of GaN on GaAs (100) at a substrate temperature of 600 °C. The spectrum in the front corresponds to the clean substrate and displays only a peak at 281 cm^{-1} caused by scattering of the longitudinal optical (LO) phonon mode of the GaAs substrate. As growth progress additional features are observed in the spectral region 500–800 cm^{-1} , which are attributed to Raman scattering in the growing layer. For a closer view at the frequency positions Raman spectra from different stages of growth are displayed in Fig. 2 corresponding to clean surface, 30, and 230 nm of layer thickness, respectively (the thickness determination will be described later). The two main peaks are located at 560 and 718 cm^{-1} , respectively. The GaN layer is likely to be composed of a hexagonal as well as a cubic phase, therefore the peak at 560 cm^{-1} is attributed to the E_2 mode of hexagonal GaN (570 cm^{-1} at RT) while the other peak arises from scattering by the A_1 mode of hexagonal GaN (735 cm^{-1} at RT) and the LO mode of cubic GaN (740 cm^{-1} at RT).¹¹ Due to the high temperature the two modes are broadened so that they cannot be resolved, therefore only a single structure is observed. This assignment is further discussed below in comparison to the results for GaN growth on Si (111). Comparing the position of the E_2 mode measured here at 600 °C with the value reported for room temperature provides an estimate of 1.7 $\text{cm}^{-1}/100$ K for the temperature induced linear phonon shift in GaN. An additional peak is observed in the uppermost spectrum of Fig. 2 at 143 cm^{-1} , which corresponds to the position of the low-frequency E_2 mode of hexagonal GaN at room temperature. The fact that this mode is fairly strong in intensity and almost unshifted compared to its room temperature position is in agreement with the observations by Perlin *et al.*¹² for this mode under pressure.

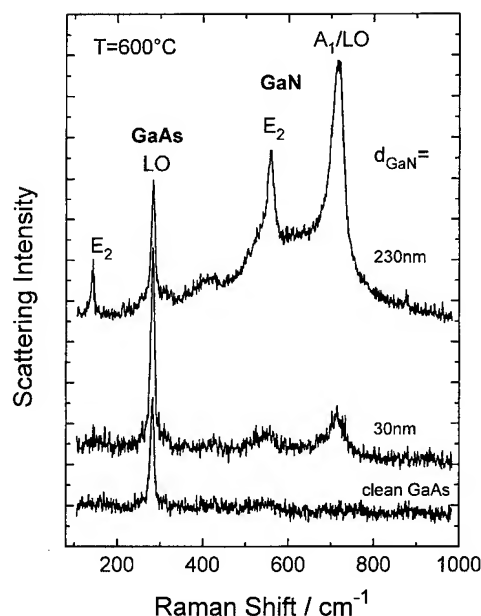


FIG. 2. Raman spectra from different stages of the growth of GaN on GaAs (100) at 600 °C.

In the region between the LO mode of GaAs and the E_2 mode of GaN additional weak structures are visible. These can be attributed to second-order Raman scattering processes of acoustic phonons that have been observed in this spectral region.¹³

As can be seen from Fig. 2 the two main GaN features are already visible after deposition of only 30 nm of GaN. The high sensitivity necessary for the detection of such a thin layer is a consequence of the resonant Raman scattering condition established by the choice of the 3.05 eV line for excitation at this growth temperature. The energies of the E_0 gaps of cubic and hexagonal GaN at RT are 3.23 eV and 3.4 eV, respectively.¹⁴ A shift of -5.4×10^{-4} eV/K has been reported for cubic GaN above room temperature,¹⁴ so the excitation energy is very close to the E_0 gap of cubic GaN at 600 °C and at least in the vicinity of the hexagonal gap. Furthermore, the proximity to the band gap is supported by the appearance of photoluminescence, which is observed as broad, increasing background in the Raman spectra. Since the development of the background is clearly correlated with the growth of the GaN layer it cannot be attributed to black body radiation at the high temperature, which should already be observable from the clean substrate.

The intensities of all features in the Raman spectra displayed in Fig. 1 show a significant modulation upon deposition time. This modulation is caused by Fabry–Perot interference of the incident as well as the scattered light within the heterostructure, and its occurrence indicates an optically smooth interface and surface.¹⁵ Furthermore, diffuse elastic light scattering that is typical for rough surfaces and would cause a strong increase of scattering intensity at the low energy side of the Raman spectra is not observed.

Using the optical constants of the materials the evolution of the intensities of both the substrate and the layer related

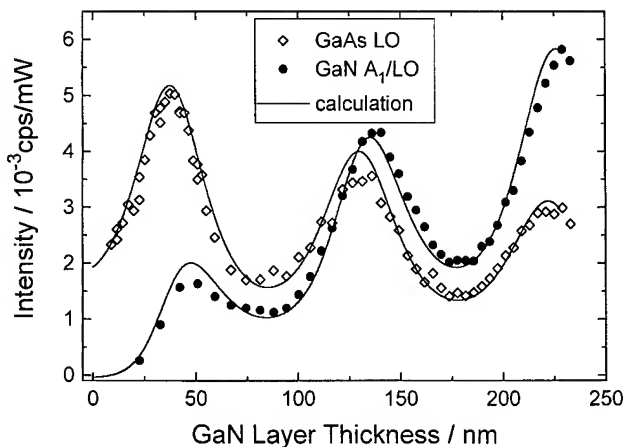


FIG. 3. Intensity of the LO mode of GaAs (open symbols) and A_1 /LO structure of GaN (solid symbols) as a function of layer thickness together with the calculated intensity evolution of substrate and layer related Raman scattering (line).

features can be calculated as a function of layer thickness. For the calculation the substrate and the surrounding vacuum are assumed semi-infinite with a plane layer of growing thickness on the substrate surface. All interfaces are assumed to be ideally smooth. Details on the model can be found elsewhere.¹⁶ By comparing the result with the observed evolution in Fig. 1 the deposition time can be correlated with the layer thickness. As a result a growth rate of 0.46 nm/min is obtained for the GaN. Figure 3 shows the observed scattering intensities of the LO mode of GaAs and the A_1 /LO structure of GaN together with the result of the calculation. A value of $n=2.2$ was estimated for the real part of the complex index of refraction $N=n+ik$ of GaN at 600 °C based on the data given in Ref. 17, while $n=3.8$ and $k=1.8$ for the GaAs substrate were extrapolated from Ref. 18. The imaginary part $k=\text{Im}(N)$ of GaN was introduced as a fit parameter, since it varies strongly in the vicinity of the E_0 gap. Best agreement is found for $k=0.035$ consistently for substrate and layer related scattering.

Alternatively GaN was grown on the (111) surface of a silicon substrate and the growth was again monitored by Raman spectroscopy. The result is displayed in Fig. 4. The first spectrum corresponding to the clean Si surface is governed by a peak at 507 cm^{-1} caused by scattering by the phonon mode of Si. The additionally broad, weak structure in the spectral region 900–1000 cm^{-1} is attributed to second order scattering processes of Si. First evidence of the formation of a GaN layer is the familiar intensity modulation of the Si peak that starts immediately after beginning the deposition. After a few minutes new features appear in the spectra that are directly related to Raman scattering in the growing layer. It is noteworthy that no photoluminescence background is observed in contrast to the results for growth of GaN on GaAs (100) (see Fig. 1). Its absence is compatible with the higher band gap energy of hexagonal GaN predominantly grown on Si (111) thus corroborating the presence of the cubic GaN phase on GaAs (100) and thereby supporting the

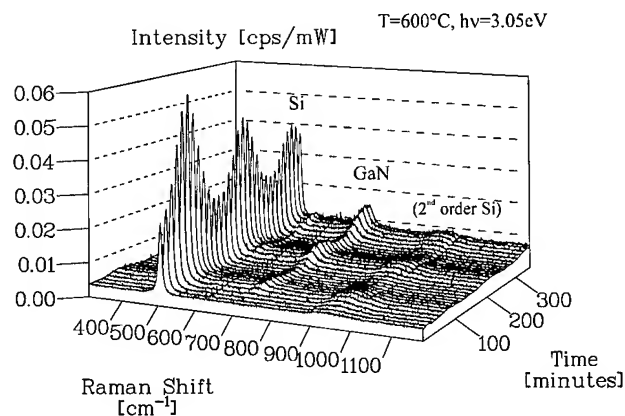


FIG. 4. Evolution of Raman spectra during growth of GaN on Si (111) at 600 °C ($h\nu=3.05$ eV).

assignment given above for the structure in the Raman spectra at 718 cm^{-1} .

Figure 5 shows the relevant part of the spectrum recorded at the end of GaN growth on the Si (111) substrate together with the spectrum obtained from growth on the GaAs (100) substrate for the same thickness of 230 nm^{-1} . The thickness scale for the growth on Si (111) was derived by the same procedure as described above assuming $N=5.6+3.0i$ for the index of refraction of the silicon substrate.¹⁹ The E_2 and A_1 /LO structures of GaN are observed in both spectra. In the case of the Si (111) substrate the A_1 /LO structure has a smaller full width at half-maximum (FWHM) (32 cm^{-1}), while its peak position is shifted slightly (1 cm^{-1}) to higher frequencies. The smaller FWHM indicates a lower degree of mixed phase disorder, which is consistent with the fact that the predominant growth of one phase, i.e., the hexagonal

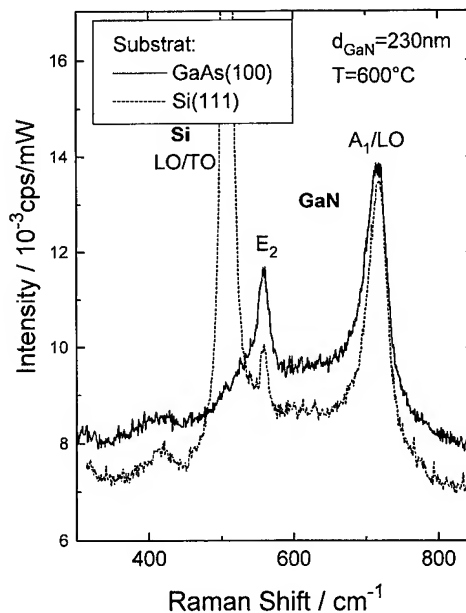


FIG. 5. Raman spectra recorded during growth at a layer thickness of 230 nm of GaN on GaAs (100) (solid line) and Si (111) (dashed line).

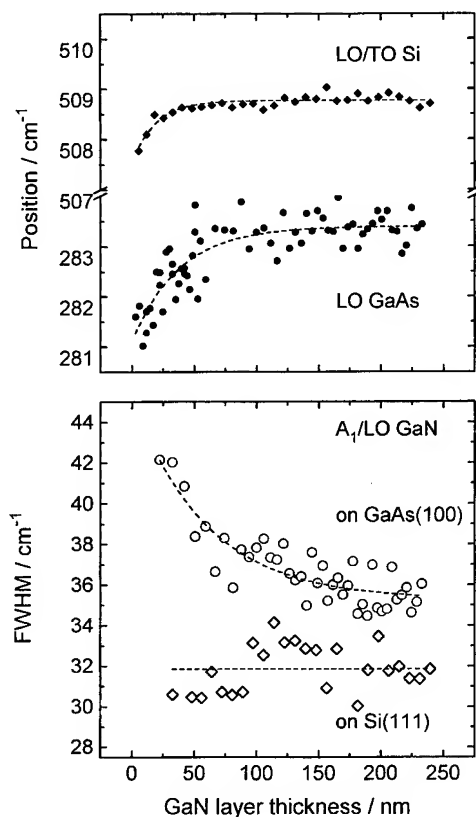


FIG. 6. Upper frame: Positions of the phonon modes of the substrate materials GaAs (solid circles) and Si (solid diamonds) as a function of GaN layer thickness. Lower frame: FWHM of the A_1 /LO structure of the GaN layer as a function of layer thickness on GaAs (100) (open circles) and Si (111) (open diamonds). The dashed lines are a guide for the eye.

modification is preferred on a (111) surface. The small increase of the peak position might be explained by the lower lattice mismatch of Si to GaN that leads to lower residual tensile strain in a layer grown on Si than on GaAs.

For a more detailed analysis of the evolution of the frequency position and FWHM during growth the phonon modes of GaAs and Si and the A_1 /LO structure of GaN were fitted with asymmetric Lorentzian line shapes after subtraction of a linear background. The asymmetry was such that the FWHM of the lower energy side of the peak (Γ_1) was larger than on the higher energy side (Γ_r). In case of the substrate phonon modes only small asymmetries are found ($\Gamma_1/\Gamma_r \approx 1.4$ for GaAs, $\Gamma_1/\Gamma_r \approx 1.2$ for Si) and are probably caused by anharmonic contributions to the scattering process introduced due to the high temperature. The line shape of the GaN structure is additionally affected by disorder resulting from the mixture of cubic and hexagonal phases, leading to ratios of Γ_1/Γ_r up to 2.2.

The upper part of Fig. 6 shows the evolution of the positions of the substrate phonon modes during the growth of GaN at 600 °C. On GaAs a shift of about 2.2 cm⁻¹ to higher frequencies is clearly observed in the initial phase of the growth process, i.e., the deposition of the first 50 nm of GaN. If this shift would be attributed to a variation of substrate temperature due to the change of emissivity it would indicate

an extremely unlikely decrease of the substrate temperature of 120°. In fact, the growth is started by opening the shutters of both the Ga and the nitrogen source, thereby exposing the substrate surface to an additional thermal radiation load which should rather result in an increase of temperature. Therefore the observed frequency shift is attributed to the development of a compressive stress of about 2 GPa in the substrate at the interface.²⁰ It should be mentioned, that the high absorption for the excitation energy of 3.05 eV leads to a very low penetration depth of 18 nm into the GaAs substrate while the GaN overlayer is comparably transparent. Consequently, the information depth of Raman scattering in GaAs is about 9 nm, thus the experiment is extremely sensitive to changes in the layer/substrate interface region. Still the stress evaluated here is surprisingly high and incompatible with the usual picture of epitaxial growth which assumes that strain relaxation via defect formation in the layer should occur in the very early stages of growth due to the extremely large lattice mismatch. On the other hand, the evolution of strain in the substrate is clearly associated with the formation of the GaN layer and cannot be assigned to, e.g., a substrate modification like nitrogen incorporation in the GaAs lattice as judged from an additional nitridation experiment.

Considering the nucleation of GaN on the Si (111) surface a similar effect as for the GaAs substrate is observed. Here the frequency of the Si phonon mode increases by about 1 cm⁻¹ during the deposition of the first 50 nm of GaN, indicating the development of a compressive stress of about 0.2 GPa at the GaN/Si interface.²¹ The enormous difference in substrate strain for GaAs and Si cannot simply be explained by the difference in lattice mismatch, but is also likely to be a result of the different substrate orientation influencing the initial nucleation of GaN. At this stage a conclusive microscopic model in order to explain the observed substrate stress effects is still missing, but this question is currently addressed by investigating the influence of different substrate surface modifications prior to the growth, e.g., nitridation or starting GaN nucleation at a lower temperature.

The lower part of Fig. 6 shows the development of the FWHM [$\Gamma = (\Gamma_1 + \Gamma_r)/2$] of the GaN A_1 /LO feature with increasing layer thickness on both substrates. Generally the linewidth of the phonon structure of GaN grown on the Si (111) surface is smaller than for growth on GaAs (100). As already mentioned this is attributed to the lower degree of mixed phase disorder in the GaN on Si, which is predominantly hexagonal. No significant dependence of the GaN phonon mode linewidth on the layer thickness is observed. In contrast to this, an overall decrease of the FWHM is found during growth on GaAs (100). This can be explained by assuming stronger competition between the hexagonal and cubic phase of GaN when grown on GaAs, especially in the initial phase of growth. As a consequence, the formation of many small domains leads to a high degree of lattice disorder which is gradually reduced when larger domains are grown with increasing layer thickness.

IV. CONCLUSIONS

Raman spectroscopy has been applied for the first time for monitoring the MBE growth of GaN on GaAs (100) and Si (111) at a temperature of 600 °C. RRS conditions were utilized in order to achieve the necessary sensitivity of the detection of ultrathin GaN layers. A significant intensity modulation is observed in the Raman spectra which allows layer thickness to be calculated. A shift of the phonon frequencies of both the GaAs and the Si substrate to higher energies in the initial phase of the growth process was found indicating the formation of compressive strain in the substrate interface region. In the case of GaAs as substrate material an improvement of crystal quality of the GaN layer as judged from the decrease of the linewidth of the GaN phonon features with increasing layer thickness was found.

ACKNOWLEDGMENTS

The authors would like to thank Dave Westwood from the University of Wales College of Cardiff for the preparation of the As-capped GaAs substrates. Financial support by the Deutsche Forschungsgemeinschaft (Grant No. DFG Za 146/7-1) is gratefully acknowledged.

¹S. N. Mohammad, A. A. Salvador, and H. Morkoc, *Proc. IEEE* **83**, 1306 (1995).

²S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto, *Jpn. J. Appl. Phys.* **1** **35**, L74 (1996).

³S. Nakamura, Y. Harada, and M. Seno, *Appl. Phys. Lett.* **58**, 2021 (1991).

⁴S. Strite, J. Ruan, Z. Li, A. Salvador, H. Chen, D. J. Smith, W. J. Choyke,

and M. Morkoc, *J. Vac. Sci. Technol. B* **9**, 1924 (1991).

⁵W. Richter and J.-T. Zettler, *Appl. Surf. Sci.* **100/101**, 465 (19916).

⁶D. Drews, M. Langer, W. Richter, and D. R. T. Zahn, *Phys. Status Solidi A* **145**, 491 (1994).

⁷D. Drews, A. Schneider, K. Horn, and D. R. T. Zahn, *J. Cryst. Growth* **159**, 152 (1996).

⁸G. Bauer and W. Richter, *Optical Characterization of Epitaxial Semiconductor Layers*, edited by G. Bauer and W. Richter (Springer, Berlin, 1976).

⁹W. Richter, *Resonant Raman Scattering in Semiconductors*, in Springer Tracts in Modern Physics (Springer, Berlin, 1976).

¹⁰V. Wagner, D. Drews, N. Esser, D. R. T. Zahn, and W. Richter, *J. Appl. Phys.* **75**, 7330 (1994).

¹¹H. Siegle, L. Eckey, A. Hoffmann, C. Thomse, B. K. Meyer, D. Schikora, M. Hankeln, and K. Lischka, *Solid State Commun.* **96**, 943 (1995).

¹²P. Perlin, C. Jaubertie-Carillon, J. P. Itie, A. San Miguel, I. Grzegory, and A. Polian, *Phys. Rev. B* **45**, 83 (1992).

¹³H. Siegle, L. Filippidis, G. Kaczmarczyk, A. Litvinchuk, A. Hoffmann, and C. Thomsen, *Proceedings of the 23rd International Conference on the Physics of Semiconductors* (World Scientific, Singapore, 1996), Vol. 1, p. 538.

¹⁴G. Ramirez-Flores, H. Navarro-Contreras, A. Lastras-Martínez, R. C. Powell, and J. E. Greene, *Phys. Rev. B* **50**, 8433 (1994).

¹⁵D. Drews, M. Langer, W. Richter, and D. R. T. Zahn, *Phys. Status Solidi A* **145**, 491 (1994).

¹⁶M. Ramsteiner, C. Wild, and J. Wagner, *Appl. Opt.* **28**, 4017 (1989).

¹⁷S. Logothetidis, J. Petalas, M. Cardona, and T. D. Moustakas, *Phys. Rev. B* **50**, 18 017 (1994).

¹⁸P. Lautenschlager, M. Carriga, S. Logothetidis, and M. Cardona, *Phys. Rev. B* **35**, 9174 (1987).

¹⁹D. E. Aspnes and A. A. Studna, *Phys. Rev. B* **27**, 985 (1983).

²⁰R. Trommer, E. Anastassakis, and M. Cardona, *Light Scattering in Solids*, edited by M. Balkanski, R. C. C. Leite, and S. Porto (Flammarion, Paris, 1976), p. 396.

²¹B. A. Weinstein and G. J. Piermarini, *Phys. Rev. B* **12**, 1172 (1975).

Nitridation of the GaAs (001) surface using atomic nitrogen

P. Hill,^{a)} D. I. Westwood, L. Haworth, J. Lu, and J. E. Macdonald

Department of Physics and Astronomy, University of Wales Cardiff, Cardiff CF2 3YB, United Kingdom

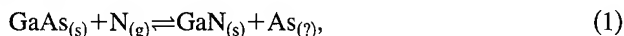
(Received 13 January 1997; accepted 5 May 1997)

The effect of active nitrogen, generated by a radio frequency plasma source, on clean GaAs (001) surfaces has been examined using x-ray photoemission spectroscopy (XPS) and reflection high energy electron diffraction. The "nitridation" of the surface was performed under fixed plasma conditions, compatible with the molecular beam epitaxial growth of GaN, and as a function of both temperature (in the range ~300–600 °C) and time (up to 15 min). At low temperatures the nitridation proceeds very slowly and was characterized, in its initial stages, by the transformation of the (2×4) reconstructed GaAs surface to a high intensity "amorphous" haze, presumed to be related to the As released in the anion exchange reaction but not evaporated from the surface. At high temperatures the nitridation is much more aggressive readily forming thicker GaN films of a polycrystalline nature. Curve fitting of the XPS spectra, to reveal the nature of the reaction products indicated the probable formation of As–N species in addition to GaN. © 1997 American Vacuum Society. [S0734-211X(97)12004-2]

I. INTRODUCTION

Recent progress in the epitaxial growth of group III nitride materials, culminating in the fabrication of high brightness light emitting diodes (LED) with wavelengths extending from the violet to the green and most recently in the demonstration of laser action at violet wavelengths,¹ has prompted a huge worldwide research and development interest. Additionally, the relatively high bond strength of the material holds the promise of future applications in high temperature electronics.² The major perceived stumbling block to the continued development of these materials is the absence of a convenient large area, lattice matched substrate. Until the work that led to the demonstration of a working LED, grown on a sapphire substrate, this problem was thought to be effectively insurmountable. Consequently, a range of different substrates have been investigated.³

One combination that has proved to be very interesting is GaN/GaAs (001)^{4–6} where, by suitable control of the growth conditions, it has been found possible to produce GaN in either its cubic or its more natural wurtzite phase. The cubic phase has been pursued since it offers a number of potential advantages over the hexagonal phase and recent reports, of optical gain in pumped cubic GaN at room temperature from samples that are featureless under an optical microscope,⁷ indicate the possibility of high quality material. Under molecular beam epitaxy (MBE) conditions using N atoms/ions as the active N species, the purity and quality of the cubic phase grown on GaAs (001) is strongly dependent on both the surface stoichiometry⁴ and the procedures adopted at the initiation of growth.⁵ Nitridation of the substrate, i.e., exposure of the surface to the active nitrogen species, has been identified as critically important. In the case of atomic N the nitridation process proceeds simply via an anion exchange reaction:⁶



^{a)}Electronic mail: Spxph@cf.ac.uk

where the uncertainty in the fate of the elemental As reflects the fact that it could remain as a solid or be released as a vapor or, as has been shown in reactions using hydrazoic acid (HN₃) as a precursor, even into As–N species.⁸ Short nitridation times subsequently resulted in cubic GaN whilst longer times result in wurtzite material, presumed to be due either to the formation of wurtzite GaN or (111) facets which act as a template for further wurtzite growth.⁵ Conversely, for metal organic vapor phase epitaxy based growth, using dimethylhydrazine (DMHy) or ammonia as a source of active nitrogen, nitridation of the GaAs surface for long periods prior to growth still results in cubic GaN material.⁹

Clearly, the effect of nitridation on the initial GaAs surface plays a crucial role in the epitaxial process. In order to understand this interface more fully we have performed a detailed x-ray photoemission spectroscopy (XPS) and reflection high energy electron diffraction (RHEED) study of the reactions between the GaAs surface and atomic nitrogen as a function of both atomic nitrogen exposure time and substrate temperature.

II. EXPERIMENT

All experiments were performed in a home-built MBE style growth chamber attached to a conventional VG ESCALAB Mk II. The system was specifically designed to allow the surface science investigation of the crucial initial stages of growth whilst maintaining the degree of control and reproducibility inherent in dedicated growth systems.

In more detail, the growth chamber is based on a six way "cross" one side of which is attached to the ESCALAB via a gate valve whilst the other ports lead to; the sample manipulator, the source and shutter flange, the pumping port (leading to a diffusion pump), a 5 kV RHEED gun, and its associated RHEED screen. The source flange holds up to three thermal evaporation sources, a pyrometer viewport,

and, most importantly for these studies, an Oxford Applied Research CARS 25 activated rf plasma source used for the production of either atomic nitrogen or hydrogen. Similar active nitrogen sources are used in many group III-nitride growth laboratories worldwide.^{10,11} All of the sources are individually shuttered although these are not intimate with the sources and some atomic nitrogen does reach the sample, as will become apparent.

All experiments were performed with fixed plasma conditions and therefore with a fixed active nitrogen flux onto the substrate. For atomic hydrogen cleaning, hydrogen gas was admitted to the plasma source via a leak valve and was operated at a chamber pressure of $(5.5 \pm 0.5) \times 10^{-5}$ mb, as measured by an ion gauge mounted into the liquid nitrogen cold trap of the system's diffusion pump. With a forward rf power of 400 W an optical emission detector (OED), based on an optical fiber fed into a Si photodiode, gave a reading of ~ 1.2 V. For nitridation, nitrogen gas admission was controlled either by a mass flow controller (mfc) or by a leak valve calibrated against the mfc. The N flow rate was ~ 1.0 sccm, corresponding to a system pressure of $(5.5 \pm 0.5) \times 10^{-5}$ mb. With a forward rf power of 450 W the OED gave a reading of (3.35 ± 0.05) V. The etched GaAs substrates, of dimension $\sim 1 \times 1$ cm, were In mounted onto Mo spades compatible with transportation by modified VG ESCA "stubs." The sample temperature was calibrated against a thermocouple (situated between the heater and the spade) by means of an Ircan V-series pyrometer and extrapolated for low sample temperatures.

XPS measurements were taken at normal emission, with an analyzer pass energy of 5 eV and using a Mg anode to give a photon energy of 1253.6 eV. Spectra were obtained for the Ga 3d, As 3d, and N 1s regions whose emitted electrons have escape depths of ~ 26 , ~ 26 and ~ 20 Å, respectively.¹²

The substrates were cleaned under an atomic hydrogen flux for 5 min at 400 °C after which time RHEED showed a (2×4) reconstruction and XPS showed the absence of any O or C contamination. Nitridation was performed, following adjustment of the sample temperature where necessary, for times up to 15 min in steps of 5 min duration. Between nitridations, at temperatures in the range 300–600 °C, the samples were transferred to the ESCALAB for measurement by XPS. Although both the Ga and As 3d and 2p core level features were measured most of the useful information has been obtained by consideration of the evolution of the Ga 3d and As 3d features in combination with the N 1s level.

III. RESULTS

A. RHEED observations

Progression of the RHEED pattern for successive nitridations of a clean GaAs (001) surface at low temperatures (300–530 °C) took the following course: For these samples the starting GaAs reconstruction was (2×4) , the same as that obtained following atomic hydrogen cleaning at 400 °C. Once the nitrogen plasma is struck, even with the plasma source shutter closed, the GaAs reconstruction features fade,

the intensity being transferred to the background, indicating that a highly disordered layer is formed on the surface. On opening the shutter an unreconstructed streaked pattern returned, the intensity of which increased slowly with time. Further nitridation results in modulation of these streaks (i.e., bulk diffraction spots begin to appear) indicative of some roughening of the surface. The rate at which modulation of the streaks occurred depended on the sample temperature, so that, for a total nitridation time of 15 min the 450 °C sample showed only slight modulation whilst for 530 °C there was little streaking left between the bulk spots.

At higher sample temperatures the progression was quite different. For example at 580 °C the following observations were made. On heating the cleaned sample, following cleaning, from 400 to 580 °C the $2 \times$ reconstruction of the (2×4) pattern changes to a $4 \times$. It was not possible to access the orthogonal direction by RHEED with this sample but it was believed, from previous samples, that the reconstruction was probably (4×6) . On striking the nitrogen plasma (with the shutter closed) the $4 \times$ reconstruction quickly changed to an excellent $3 \times$, consistent with previous reports of a (3×3) reconstruction occurring at the early stages of nitridation at this temperature.^{13,14} After a further ~ 10 s the $3 \times$ and bulk diffraction features faded, again indicative of a highly disordered layer on the surface. Later still sharp bulk spots began to appear, although on opening the plasma shutter these quickly gave way to a ringed, or polycrystalline, pattern which, once established, changed little.

B. Photoemission results: Ga 3d core level peak shapes

As the temperature and time of nitridation increased the Ga 3d feature increased in intensity, broadened, and slowly shifted in position to higher binding energy (BE). This was consistent with the Ga 3d feature being formed by the superposition of contributions from the original GaAs substrate and from the GaN film above it. The Ga–N related features have a higher BE than the Ga–As features, as has been previously reported.⁶

In order to confirm the above view and extract more information regarding the nitridation process the Ga 3d spectra from the clean GaAs surface, the nitrided surfaces and a thick GaN film, grown *in situ* onto a GaAs (001) sample, were curve fitted. Although the Ga 3d core level is in reality a doublet, as a result of spin orbit splitting, it was found that, following removal of a linearly sloping background, the clean GaAs feature could be well represented by a single Voigt function with Lorentzian and Gaussian widths of 0.30 and 0.81 eV, respectively. With this as a guide the following scheme was used to fit the reacted Ga features: (a) the background was assumed to be linear in all cases; (b) following nitridation the Ga 3d peak was assumed to be composed of only two Voigt functions; (c) the Lorentzian widths of both peaks (GaAs and GaN) were constrained to the clean GaAs value; (d) both peaks were initially assumed to have Gaussian widths of 0.81 eV, if a good fit could not be obtained with this width the GaN Gaussian width was allowed to

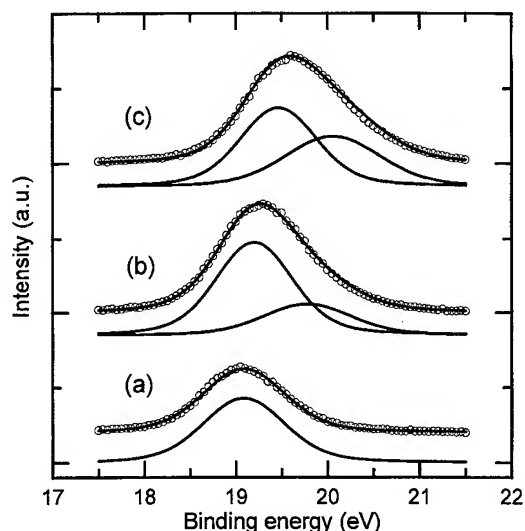


FIG. 1. XPS scans of the Ga 3d spectral regions for (a) the clean GaAs surface, (b) following 5 min nitridation at 300 °C, (c) following 5 min nitridation at 450 °C. The background has been subtracted, the fit to the curve passes through the data points, and the individual components are shifted down slightly for clarity.

broaden; (e) the GaN features were allowed to vary in binding energy—it was not usually necessary to let the GaAs peak position vary. Using this scheme good fits to Ga 3d spectra were obtained. This is demonstrated in Fig. 1 which shows, in addition to the original clean GaAs Ga 3d signal, curve fitted spectra obtained following 5 min nitridation. The chemical shifts between the two Ga 3d contributions being in the range 0.56–0.63 eV agree well with the report of DeLouise⁶ following bombardment of the GaAs (110) surface with N_2^+ .

C. Photoemission results: As 3d core level peak shapes

The changes in the As 3d core level appear most dramatic at the early stages of nitridation at low temperature, where a reacted As species is observed quite distinct from the substrate contribution. These features were curve fitted using a similar scheme as for the Ga 3d. However, in this case the larger spin orbit splitting of the As 3d level meant that the clean surface was best fitted with a doublet of splitting 0.67 eV, branching ratio 1.5 and using Gaussian and Lorentzian widths of 0.63 and 0.42 eV, respectively. Due to the fact that the reacted As components decreased, relative to the substrate contribution, with sample temperature all As 3d spectra could be sensibly fitted. Figure 2 shows, in addition to the original clean GaAs As 3d signal, curve fitted spectra obtained following 5 min nitridation at 300 and 450 °C. The separation of the two contributions is in the range 1.4–1.85 eV. At higher nitridation temperatures only the substrate feature remained, although with considerably reduced intensity. Thus the As 3d feature provides the best indication of the attenuation of the substrate signal, and therefore the thickness of the GaN, at these temperatures.

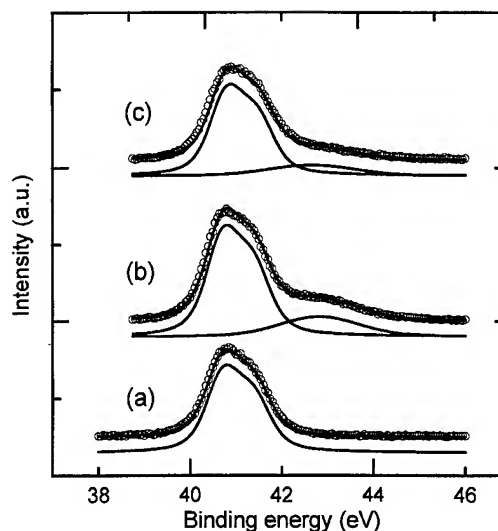


FIG. 2. XPS scans of the As 3d spectral regions for (a) the clean GaAs surface, (b) following 5 min nitridation at 300 °C, (c) following 5 min nitridation at 450 °C. The background has been subtracted, the fit to the curve passes through the data points, and the individual components are shifted down slightly for clarity.

Combining the evidence from the low nitridation temperature curve fitted Ga 3d features and the As 3d attenuation at high temperatures it can be stated that the thickness of the GaN increased with both nitridation time and temperature and, in particular, that no upper thickness limit to the nitridation process was evident.

D. Photoemission results: N 1s core level peaks

For all experiments performed the N 1s feature appeared as a symmetric single peak. However, examination of the full width at half-maximum (FWHM) did provide some evidence that nitrogen exists in more than one chemical state. The weakest N 1s signal was, as expected, obtained for the shortest time, lowest temperature nitridation, i.e., for 5 min at 300 °C and had a FWHM of 1.39 eV. The highest FWHM, of 1.58 ± 0.5 eV, was obtained following subsequent nitridations at 300 °C and for nitridations of 5–15 min at 450 °C. By contrast, at higher temperatures, where the extent of nitridation was more advanced, no reacted As species were detected and the N 1s FWHM reduced to 1.30 eV, the same value as that obtained for thick epitaxially grown GaN. It can be speculated that the N 1s broadening is due to As–N species.

E. Photoemission results: Intensity and attenuation data

It is common, when analysing XPS data of thin film growth, to plot measured intensities against some process parameter, in our case nitridation time or temperature. Unfortunately, the process of performing these experiments required a large amount of sample manipulation, between the nitride growth chamber and the ESCALAB analysis chamber, resulting in uncertainties in absolute intensity measure-

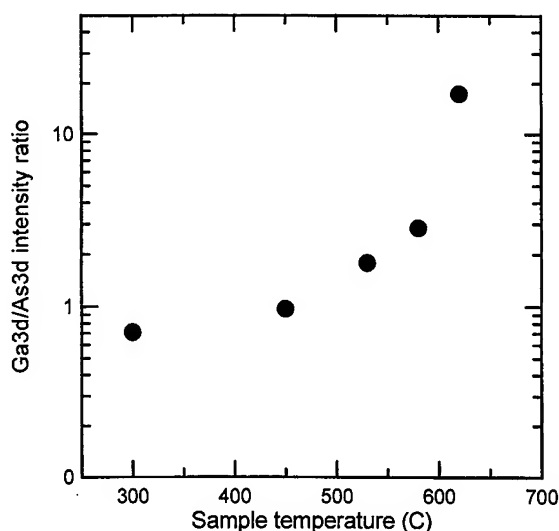


FIG. 3. Ratios of the total areas under curves (Ga 3d/As 3d), following background subtraction, as a function of nitridation temperature following 5 min nitridation using atomic nitrogen.

ments. However, comparison of relative intensities of different peaks, obtained under identical conditions without sample manipulation, should be much more reliable.

The simplest way to present such information, which avoids possible errors in curve fitting, is to display the ratios of the total areas under curves, following background subtraction. Figures 3 and 4 show the ratios for (Ga 3d)/(As 3d) and (Ga 3d)/(N 1s), respectively, as functions of nitridation temperature and 5 min nitridation. It should be noted, as has previously been discussed, that the Ga and As 3d's contain both substrate and overlayer contributions and the N 1s probably also has more than one (overlayer) contribution. Noting that the (Ga 3d)/(As 3d) ratio measured for clean GaAs was (0.65 ± 0.03) , Fig. 3 shows clearly that the extent

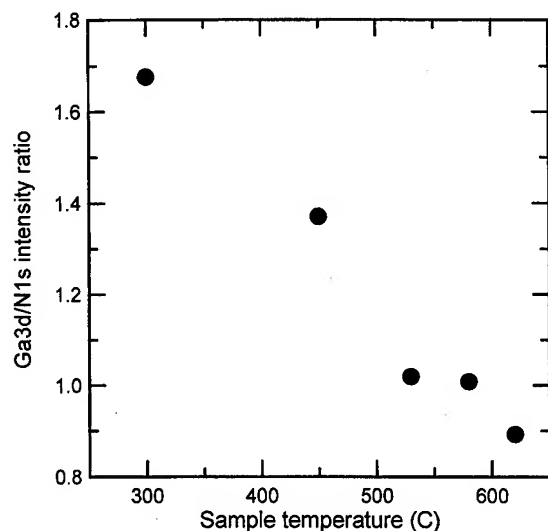


FIG. 4. Ratios of the total areas under curves (Ga 3d/N 1s), following background subtraction, as a function of nitridation temperature following 5 min nitridation using atomic nitrogen.

of nitridation (i.e., thickness of GaN) was very low at 300 °C, but increased strongly as a function of increasing temperature. Figure 4 shows that at low temperature the (Ga 3d)/(N 1s) ratio is dominated by the influence of the Ga 3d signal from the substrate, whereas at the high temperatures it is consistent with the formation of a relatively thick GaN film [for thick layers of epitaxial GaN the (Ga 3d)/(N 1s) ratio has been found to be (0.95 ± 0.10)].

To actually quantify the extent of nitridation it is necessary to calculate the areas under specific peaks, which for XPS may be obtained using the following equation:

$$I = K \int_{t_1}^{t_2} \exp\left(\frac{-l}{\lambda}\right) dl, \quad (2)$$

where I is the intensity of electrons emitted from the material at a certain kinetic energy, t_1 and t_2 are the thickness limits of the layer, l is the distance below the surface (for normal emission), λ is the electron escape depth for the particular materials and electron kinetic energy considered, and K is a constant taking into account (a) the number density of atoms in the material, (b) sample area, (c) differential photoionisation cross section, (d) X-ray flux and detector efficiency.¹⁵ By considering the relative intensities of contributions from the same element and core level (e.g., GaAs and GaN contributions to the Ga 3d signal) the integral in Eq. (2) is much simplified and the thickness of overlayers can be calculated. A consideration of the uncertainties in parameters in Eq. (2), in particular in the electron escape depth, leads to an estimated possible error of $\pm 25\%$ in the calculated thicknesses. In addition, implicit in this treatment is an assumption that the layers are all both uniform and abrupt, the consequences of which will be discussed below.

Two different approaches have been taken to quantify the extent of nitridation. The first makes use of the fact that the Ga and As 3d features could be reasonably curve fitted to reveal their substrate and overlayer contributions. Thus the relevant reacted and substrate components could be used to calculate thicknesses. This is slightly complicated by the fact that there are two reacted species present so that the composition and nature of the overlayer is unclear. Here it has been assumed, for the purposes of determining the elemental number densities, that Ga in the overlayer is in the form GaN and that the As is in the form of elemental As. To account for the unknown physical location of the "elemental As" with respect to the "GaN" a thickness range was determined by performing the calculations twice for the extreme cases of the As being located at the surface and, conversely, at the buried interface. Since such calculations are really only valid for uniform overlayers, the area ratio data and results of these calculations are presented in Table I only for the two lowest substrate temperatures (300 and 450 °C) where RHEED and *ex situ* atomic force microscopy (AFM) indicated that the surfaces were still smooth. Also presented in Table I are the area ratios of Ga 3d (GaN)/N 1s and Ga 3d(GaAs)/As 3d(GaAs). By comparison with the measured ratios for GaAs substrates and GaN epitaxial films, given earlier, it can be seen that, while the substrate data agrees

TABLE I. Intensity ratios obtained from the curve fitting of the XPS data from nitrided GaAs (001) surfaces at low temperature and the subsequently calculated estimates of the GaN and As overlayer thicknesses.

	Ga 3d (GaN/GaAs)	As 3d (As/GaAs)	d(GaN)/Å	d(As)/Å	Ga 3d (GaN/N1s)	(Ga 3d)/ (As 3d) (GaAs)
300 °C/5 min	0.38	0.26	3.4–3.8	2.7–3.1	0.47	0.64
300 °C/15 min	0.6	0.29	6.3–7.1	2.7–3.4	0.55	0.61
450 °C/5 min	0.77	0.18	8.1–8.7	1.6–2.2	0.62	0.66
450 °C/15 min	1.0	0.17	9.9–10.6	1.4–2.1	0.55	0.65

well, there is almost twice the N 1s signal required to satisfy the formation of GaN! Since the N must be bonded onto the surface it is again suggested that As–N species are present, and in comparable quantities to GaN, at these early stages of nitridation. In addition, comparison of the nitride thicknesses as a function of time indicate that the nitridation rate does saturate at these temperatures, presumably due to the requirement of transporting the incoming atomic nitrogen through the surface layer.

A second approach was also considered. Here, the fact that As 3d signals were detected in all experiments, and the (Ga 3d)/(As 3d) ratio for GaAs substrates is known, was used to determine the contribution of the substrate Ga to the total Ga 3d signal. The estimated thicknesses, d , determined in this way for 5 min nitridations at the same temperatures considered were as follows: $T = 300\text{ °C}$ ($d = 4.8\text{--}5.5\text{ Å}$) and $T = 450\text{ °C}$ ($d = 8.6\text{--}9.3\text{ Å}$). The spread of values again reflects the uncertainty in the location of the reacted As as discussed. Comparison of the derived thicknesses, between the two methods for the low temperature nitridations shows that the agreement, although not perfect, is nonetheless reasonable.

The same calculations have also been performed for the higher sample temperatures where RHEED indicated the absence of a smooth surface. For these samples Eq. (2) is not valid and so the values derived are presented only to illustrate, in more comprehensible terms than Figs. 3 and 4, the effect of sample temperature on the extent of nitridation. Examination of Eq. (2) leads to the expectation of an enhanced substrate signal from the thinner overlayer regions giving rise to calculated thicknesses that are less than the average thickness. The calculated values are as follows: $T = 530\text{ °C}$ ($d = 16\text{ Å}$); 580 °C ($d = 25\text{ Å}$); 620 °C ($d = 62\text{ Å}$).

IV. DISCUSSION/CONCLUSIONS

The measurements presented here have shown that reaction of atomic nitrogen with the clean GaAs (001) surface is a strong function of the sample temperature and at high temperatures is capable of producing reasonably thick, but non-uniform, GaN films. The calculated thickness value of 62 Å obtained for 5 min nitridation at 620 °C must be a lower limit estimate due to this nonuniformity. In fact the aggressive nature of the process under these conditions was evidenced by the fact that the surface of the sample was visibly disrupted with a large number of local defects separated by a

generally slightly roughened background. Closer examination with an AFM revealed that the background was composed of small grains of diameter $250\pm 50\text{ Å}$ clustered together to form larger grains of diameter $1750\pm 250\text{ Å}$ which combined to give the surface a rms roughness of $\sim 40\text{ Å}$ which is comparable with the calculated thickness of the film. This granular structure is presumably responsible for the appearance of polycrystalline rings in the RHEED pattern.

It is possible that the disruption of the surface is a natural consequence of the amount of GaN being produced. Sato, using plasma enhanced metalorganic chemical vapor deposition temperature of 500 °C , observed roughening of the GaAs surface for GaN coverages of $>1\text{ ML}$ and proposed a model based on the lattice mismatch between GaAs and GaN to explain the formation of subsurface GaN clusters.¹⁶ In addition, there is a serious question as to the fate of the As that must be released at the buried interface as nitridation proceeds. Given that it is unlikely that As, as a relatively large atom, can migrate away from the interface at the same rate as nitrogen atoms can migrate to it, it is possible that the release of As contributes to the disruption of the surface region.

The nature of the surface film and the fate of the As released at low temperatures is also interesting. RHEED and AFM evidence for nitridation at 450 °C showed that this film was commensurate with the GaAs substrate and had a surface that was essentially smooth despite the fact that the measured GaN film had a thickness approaching 10 Å , a value well in excess of the $\sim 2.25\text{ Å}$ value that corresponds to 1 ML . Thus it appears possible to freeze out the condition that layers above 1 ML GaN coverage form relaxed GaN clusters.¹⁶ As has been stated, the released As had a binding energy consistent with the elemental form, however, there is evidence from the apparent excess of nitrogen present and the broadening of the N 1s signal that As–N species might also be present. This is not a new concept, having been demonstrated for the reaction of hydrazoic acid (HN_3) with GaAs (001).⁸ It can be added that the reacted As species is likely to be in a very disordered state since the Gaussian widths used to fit it were observed to increase by 2.2 to 2.5 eV compared to the width (of 0.625 eV) required as part of the fit of the clean GaAs surface As 3d doublet. At the same time the reacted Ga 3d component required an increase in the width of its Gaussian component by amounts from 0.2 to 0.45 eV

above the clean GaAs value of 0.81 eV (noting that the Ga 3d, although a doublet, has been treated as a singlet).

It is important to consider the results presented here in terms of the growth of GaN/GaAs structures. Most growth of nitrides is initiated at temperatures of $\sim 600^\circ\text{C}$ or higher, where the interaction with the surface is aggressive. This possibly goes some way to explaining the relatively poor quality of material often reported using GaAs substrates. Avoiding surface nitridation is not necessarily straightforward. As has been noted by Sato,¹⁶ plasma sources require a short but not insignificant period of time to achieve a stable flux of the active species as a result of either the transients involved in striking the plasma, or the requirement of altering the tuning conditions for coupling the rf power input. Additionally, shuttering of the plasma sources in the MBE environment is unlikely ever to be 100% efficient. Our experience with a simple shutter arrangement has been that nitridation of the clean GaAs surface occurs to some extent even with the shutter closed. Since nitridation at higher temperatures appears to be so rapid, lack of attention to the first stages of the growth process may result in poor quality or irreproducible epitaxy. Curiously, Hausenstein *et al.*¹³ have shown that the formation of the GaAs-N (3×3) reconstruction was a reversible process so that it was possible to recover a clean high quality GaAs surface especially if an As flux is also present. The implication here is that reevaporation of N occurs from this reconstructed surface so that if the active N flux is low compared to the reevaporation rate (e.g., achieved by having a very efficient shutter) high quality surfaces may be retained for long periods. The issue is clearly complicated.

In summary, the nitridation of the GaAs (001) surface with atomic nitrogen proceeds slowly at low temperatures but at the temperatures typical of epitaxial growth (600°C and higher) it is very aggressive. Therefore great care must be taken to control this process before the onset of growth. Curve fitting of the XPS spectra, to reveal the nature of the reaction products at low temperatures, indicates the possible

formation of As-N species in addition to GaN and yields some quantitative information on the thickness of the nitride layer. In view of the uncertainty in the physical location of the As for low nitridation temperatures and the nonuniformity at higher temperatures, these deduced thicknesses should be viewed as an approximate guide.

ACKNOWLEDGMENTS

The authors would like to thank the technical staff at UWC for their efforts in the design and construction of the nitride growth chamber, in particular, Chris Dunscombe, Graham Jones, Hugh Lang, Dave Pickering, Glyn Summers, and Jeff Trivett. The financial support of the UK EPSRC is gratefully acknowledged and we would also like to thank Dr. K. Narain for performing AFM measurements using a Digital Instruments Multimode system.

- ¹S. Nakamura, *Proceedings of the International Conference on the Physics of Semiconductors* (World Scientific, Singapore, 1996), Vol. 1, p. 11.
- ²M. A. Kahn, A. R. Bhattarai, J. N. Kuznia, and D. T. Olsen, *Appl. Phys. Lett.* **63**, 1214 (1993).
- ³S. Strite and H. Morkoç, *J. Vac. Sci. Technol. B* **10**, 1237 (1992).
- ⁴H. Yang, O. Brandt, and K. Ploog, *J. Electron. Mater.* **25**, 787 (1996).
- ⁵A. Kikuchi, H. Hoshi, and K. Kishino, *Jpn. J. Appl. Phys.* **33**, 688 (1994).
- ⁶L. A. DeLouise, *J. Vac. Sci. Technol. A* **10**, 1637 (1992).
- ⁷R. Klann, O. Brandt, H. Yang, H. T. Grahm, and K. H. Ploog, *Appl. Phys. Lett.* **70**, 1076 (1997).
- ⁸Y. Bu and M. C. Yin, *Surf. Sci.* **317**, 152 (1994).
- ⁹S. Miyoshi, K. Onabe, N. Ohkouchi, H. Yaguchi, and R. Ito, *J. Cryst. Growth* **124**, 439 (1992).
- ¹⁰T. S. Cheng, L. C. Jenkins, S. E. Hooper, C. T. Foxon, J. W. Orton, and D. E. Lacklison, *Appl. Phys. Lett.* **66**, 1509 (1995).
- ¹¹D. Schikora, N. Hankeln, D. J. As, K. Lishka, T. Litz, A. Waag, T. Buhrow, and F. Henneberger, *Phys. Rev. B* **54**, R8381 (1996).
- ¹²H. Gant and W. Mönch, *Surf. Sci.* **105**, 217 (1981).
- ¹³R. J. Hausenstein, D. A. Collins, X. P. Cai, M. L. O'Steen, and T. C. McGill, *Appl. Phys. Lett.* **66**, 2861 (1995).
- ¹⁴O. Brandt, H. Yang, B. Jenichen, Y. Suzuki, L. Däweritz, and K. H. Ploog, *Phys. Rev. B* **52**, R2253 (1995).
- ¹⁵T. E. Madey, J. T. Yates, and L. E. Erickson, *Chem. Phys. Lett.* **19**, 487 (1973).
- ¹⁶M. Sato, *Jpn. J. Appl. Phys.* **34**, 1080 (1995).

Optical transitions in InGaN/AlGaIn single quantum wells

K. C. Zeng, M. Smith, J. Y. Lin, and H. X. Jiang^{a)}

Department of Physics, Kansas State University, Manhattan, Kansas 66506-2601

J. C. Robert, E. L. Piner, F. G. McIntosh, and S. M. Bedair

Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, North Carolina 27685-7911

J. Zavada

U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709-2211

(Received 12 January 1997; accepted 23 March 1997)

The optical transitions in InGaN/AlGaIn single quantum wells (SQWs) grown by metal-organic chemical vapor deposition have been studied. The spectral lineshape and the recombination dynamics of the optical transitions have been systematically investigated at different conditions. It was found that the main photoluminescence (PL) emission line in these SQW was contributed predominantly by the localized exciton recombination. However, time-resolved PL results revealed the presence of a band-to-impurity transition which cannot be resolved spectroscopically from the localized exciton transition line due to the broad emission linewidth. © 1997 American Vacuum Society. [S0734-211X(97)08604-6]

I. INTRODUCTION

The group III-nitride wide band gap semiconductors have attracted much attention recently because of their potential for many applications such as blue/UV light emitting diodes (LEDs), laser diodes (LDs), and high-temperature and high-power electronic devices.¹ In particular, InGaIn has the advantage of tunability of the energy gap, allowing greater control over the spectrum of emitted light from visible to near UV. InGaIn/AlGaIn single quantum wells (SQWs) are being used as an active medium for commercial high brightness blue/green LEDs.² Blue/UV LDs operating at room temperature based on multiple quantum wells (MQW) of In_xGa_{1-x}N/In_yGa_{1-y}N ($x > y$) have also been demonstrated recently.³ However, the optical properties of InGaIn and the nature of the optical transitions in InGaIn QWs are not well understood. Needless to say, a better understanding of optical transitions in InGaIn QWs is needed for the improvements of optoelectronic devices based on these materials. In this work, we have employed time-resolved photoluminescence (PL) measurements to study the mechanisms of optical transitions in InGaIn/AlGaIn SQW. Compared with InGaIn/GaN and GaN/AlGaIn QW structures,⁴ the InGaIn/AlGaIn QW structure provides more controlled parameters due to the fact that both the barrier and well materials are alloys, an enhanced photon confinement due to the larger difference in the indices of refraction between InGaIn and AlGaIn, and an enhanced carrier confinement due to the larger band offset.

II. SAMPLES

Unintentionally doped InGaIn/AlGaIn SQW samples used here were grown by metal organic chemical vapor deposition (MOCVD) in an atmospheric pressure vertical reactor. The structure of these InGaIn/AlGaIn SQW is shown schemati-

cally in Fig. 1. Source gases used were ammonia (NH₃), trimethylgallium (TMG, -10 °C), and trimethylaluminum (TMA, +18 °C) and ethyldimethylindium (EdMIn, +10 °C) was used as the carrier gas. Prior to the deposition of the SQW, a thin 100 Å AlN buffer layer was grown on a (0001) sapphire substrate at 700 °C by atomic layer epitaxy (ALE). Next, a 0.1-μm-thick Al_yGa_{1-y}N ($y \sim 0.05$) cladding layer was grown by MOCVD at 950 °C, followed by a graded In_xGa_{1-x}N layer deposited while the temperature was ramped down from 800 to 750 °C. The active In_xGa_{1-x}N was then grown at 750 °C, followed by a 500 Å Al_yGa_{1-y}N ($y \sim 0.05$) upper cladding layer grown at 950 °C to complete the structure. The thickness of the active In_xGa_{1-x}N well is estimated to be about 50 Å, based on the growth rate and the transmission electron microscope (TEM) studies carried out on this material system. The value of x was estimated in the 20%–25% range as determined by x-ray diffraction studies on thicker In_xGa_{1-x}N films. This range in the value of x is due to compositional nonuniformity across the wafer in our current ALE/MOCVD reactor. From these parameters, the energy band diagram of the InGaIn/AlGaIn single quantum well sample studied here can be constructed and is shown in Fig. 2.

III. EXPERIMENT

For photoluminescence measurements, samples were attached to copper sample holders and placed inside a closed-cycle He refrigerator with a temperature variation from 10 to 320 K. A temperature controller enabled us to stabilize the temperature to within 0.1 K. Photoluminescence spectra were collected in a reflecting mode. Excitation pulses of about 7 ps at a repetition rate of 9.5 MHz were provided by a cavity-dumped dye laser (Coherent 702-2CD) with Rhodamine 6G dye solution, which was pumped by an yttrium-aluminum-garnet (YAG) laser (Coherent Antares 76) with a frequency doubler. The output from the dye laser was

^{a)}Electronic mail: Jiang@Phys.ksu.edu

500 Å	$\text{Al}_{0.05}\text{Ga}_{0.95}\text{N}$
50 Å	$\text{In}_{0.5}\text{Ga}_{0.5}\text{N}$
400 Å	graded $\text{In}_x\text{Ga}_{1-x}\text{N}$ ($x \sim 0-0.3$)
1000 Å	$\text{Al}_{0.05}\text{Ga}_{0.95}\text{N}$
100 Å	AlN buffer
Sapphire	

FIG. 1. Schematic diagram of MOCVD grown InGaN/AlGaIn single quantum well samples used in this work.

frequency doubled again by a second frequency doubler to provide tunability in the UV region. The laser output after the second doubler has an average power of about 20 mW, a tunable photon energy up to 4.5 eV, and a spectral width of about 0.2 meV. The laser beam size on the sample was about 0.3 mm in diameter. A single photon counting detection system was used to record the time-resolved photoluminescence spectra. With the use of a microchannel-plate photomultiplier tube (MCP-PMT), the overall time resolution of the detection system was about 20 ps. The excitation intensity was controlled by a set of UV neutral density filters with different attenuation parameters D , and was thus proportional to 10^{-D} .

IV. RESULTS AND DISCUSSIONS

Figure 3 shows the continuous-wave (cw) photoluminescence (PL) emission spectra of an InGaN/AlGaIn SQW measured at $T=10$ K. The main peak at about 2.8 eV results from the InGaN active well region and is predominantly due to the localized exciton recombination. An additional emission line with a much lower emission intensity and a narrower linewidth at about 3.62 eV is also evident in Fig. 1, which results from optical recombination in the $\text{Al}_y\text{Ga}_{1-y}\text{N}$ barrier layers. This assignment is based on the fact that the emission energy is consistent with the expected Al composition of about 0.05.

Figure 4 shows the cw PL spectra measured at $T=10$ K for five representative excitation intensities, $I_{\text{exc}}=0.01, 0.1,$

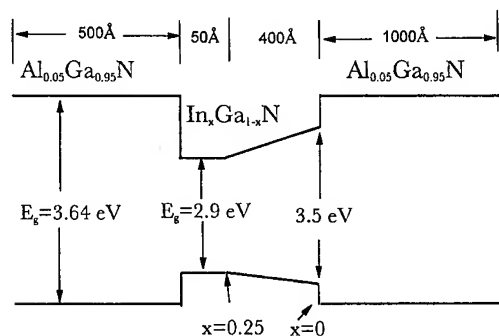


FIG. 2. Energy band diagram of InGaN/AlGaIn single quantum well sample used here.

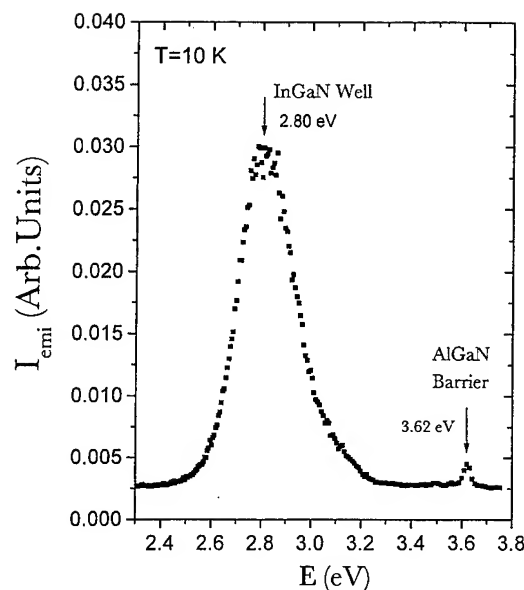


FIG. 3. PL spectrum of an InGaN/AlGaIn SQW measured at 10 K.

0.32, 0.63, and $1.0I_0$. These spectra have been normalized to the maximum peak intensity for each I_{exc} for a better illustration. The overall spectral line shapes for different I_{exc} are quite similar. However, it is clear that the peak position of the dominant band is shifted toward higher energies as excitation intensity increases. This observation is expected for a localized exciton recombination. In semiconductor alloys, excitons are localized in random potential wells induced by alloy disorder. At higher excitation intensities, lower energy

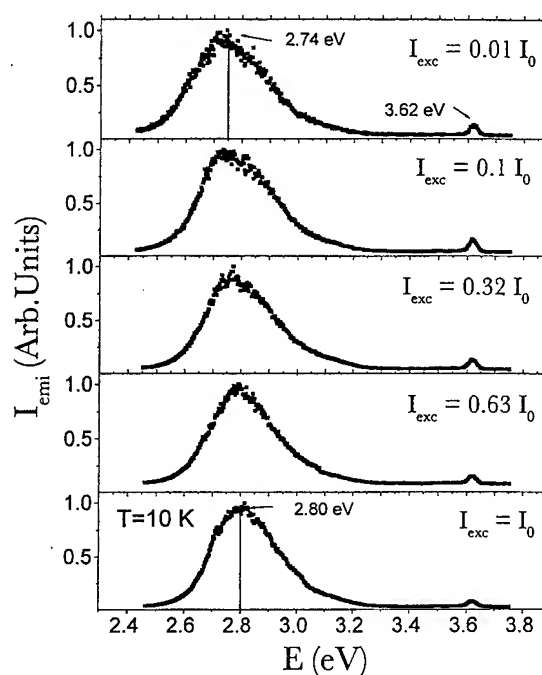


FIG. 4. PL spectra measured at $T=10$ K for five representative excitation intensities, $I_{\text{exc}}=0.01, 0.1, 0.32, 0.63,$ and $1.0I_0$, respectively.

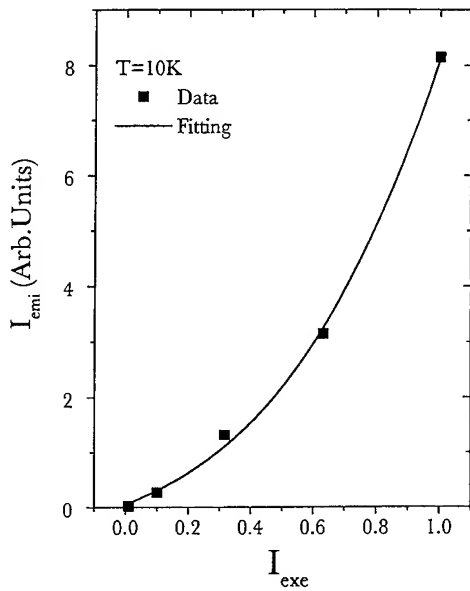


FIG. 5. Excitation intensity dependence of PL emission intensity measure at 10 K. The solid curve is the least squares fit of data by Eq. (1).

states are more likely being occupied and consequently more excitons have to fill in higher energy states, which causes the shift of emission peak position toward the higher energies.

The PL emission intensity, I_{emi} , of the dominant PL band near 2.8 eV at 10 K as a function of excitation intensity, I_{exc} , has been measured. This is plotted in Fig. 5. It is found that I_{emi} increases superlinearly with excitation intensity, I_{exc} , following a power-law form,

$$I_{\text{emi}} \propto I_{\text{exc}}^{\beta} \quad (1)$$

The solid curve in Fig. 5 is the least squares fit of the data with Eq. (1), where the fitted exponent β is about 1.4. The superlinear increase of I_{emi} with I_{exc} has also been observed for the localized exciton recombination in InGaIn epilayers,⁵ in which the exponent β is 2.6. The superlinear behavior seen here supports our interpretation that the main emission line in these InGaIn SQW is of a predominantly intrinsic nature, i.e., localized exciton recombination. The smaller value of exponent β seen here may be due to the presence of an impurity related transition which cannot be resolved spectroscopically from the localized exciton transition line due to the broad linewidth.

In order to elucidate further the nature of the dominant emission line near 2.8 eV in InGaIn/AlGaIn SQW, we have employed a time-resolved PL to study its dynamical behavior. There have been a few prior measurements on time-resolved PL on similar systems, such as InGaIn epilayers⁵ and InGaIn/GaN SQW.⁶⁻⁸ The decay of the localized excitons in high-quality InGaIn epilayers is exponential, while the PL in InGaIn/GaN SQW under a strong excitation decays according to a bimolecular law due to the nature of the band-to-band (free electron-to-free hole) transition under such a condition. On the contrary, the decay of the PL emission band in InGaIn/AlGaIn SQW observed here is neither an exponential nor a bimolecular type. This is illustrated in Fig. 6

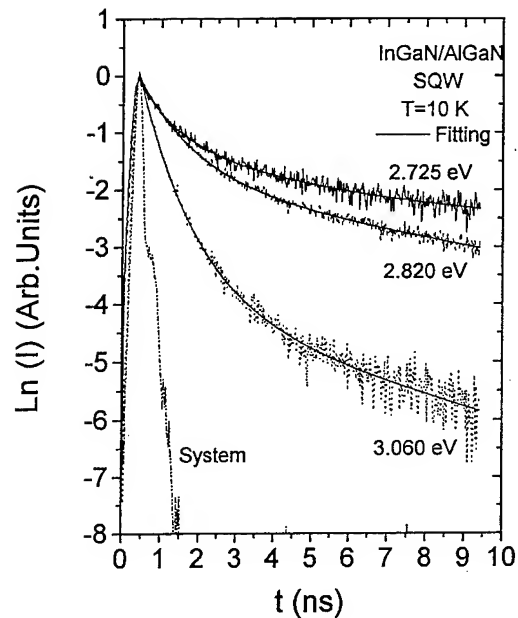


FIG. 6. Semilogarithmic plot of PL temporal responses measured at three representative emission energies. The solid curves are the least squares fit of data by two-exponential decay functions of Eq. (2). The detection system response (~ 20 ps) to the laser pulses (7 ps) is indicated as "system."

where the temporal responses of PL emission measured at 10 K for three representative emission energies together with a system response (20 ps) to the laser pulses are shown. The PL decay shown in Fig. 6 can be described very well by a two-exponential function,

$$I(t) = A_1 \exp(-t/\tau_1) + A_2 \exp(-t/\tau_2), \quad (2)$$

where τ_1 and τ_2 are the recombination lifetimes for the faster and slower components, respectively. Across the entire emission band, PL decay kinetics can all be fit very well by Eq. (2) and the PL amplitude of the fast decay component A_1 is always greater than 90%.

In Fig. 7, we plot the recombination lifetimes of the fast and slow decay components, τ_1 and τ_2 , as functions of emission energy. τ_1 is approximately constant in the range from 2.67 to 2.8 eV and then decreases from 0.7 to 0.4 ns as the emission energy increases from 2.8 to 3.05 eV and is comparable to the recombination lifetimes of the localized excitons in InGaIn epilayers.⁵ Localized excitons in semiconductor alloys can transfer from higher energy sites to lower energy sites with a relaxation rate that increases with an increase of emission energy.⁶ Thus the observed decrease of τ_1 with increasing emission energy shown in Fig. 7(a) is also expected for localized excitons. These results further support our interpretation that the main emission band in InGaIn SQW is contributed predominantly by the localized exciton recombination. However, there is a second PL emission component ($A_2 < 10\%$) which has a recombination lifetime (τ_2) ranging from 5.5 to 2.5 ns. The observed nanosecond recombination lifetimes of τ_2 suggests that this second emission component is of a band-to-impurity nature. Addition-

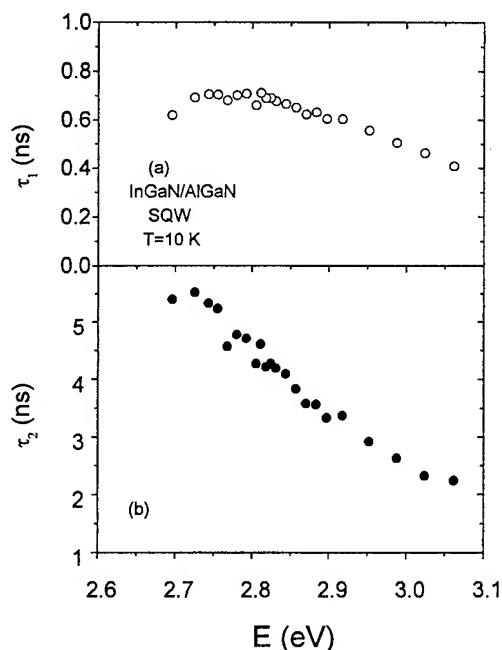


FIG. 7. The emission energy dependence of (a) the fast recombination lifetime, τ_1 , and (b) the slow recombination lifetime, τ_2 .

ally, the emission energy dependence of the recombination lifetime τ_2 is also similar to that of a band-to-impurity transition seen previously in a GaN epilayer.⁹

Furthermore, based on a previous theoretical model,¹⁰ the asymptotic decay of the band-to-impurity recombination at longer times follows a power law, $I(t) \propto t^{-\alpha}$. We have re-plotted PL temporal responses in a double logarithmic scale and found that the PL decay at longer times indeed follows a power law with the decay exponent α increasing with an increasing of the emission energy. A power law decay at longer times has been observed for a band-to-impurity recombination in *p*-type doped GaN previously.¹¹ This further supports our interpretation that the second slower decay component is due to a band-to-impurity recombination.

Finally, we have also obtained some preliminary results on recently grown InGaN/AlGaIn SQW of higher quality, in which more than one emission peak can be clearly resolved in time-resolved PL spectra. This further corroborates our assignment. The dependence of τ_2 on emission energy at low temperatures suggests that the energy level of the impurity involved in the transition may have a distribution, which together with alloy disorder in InGaIn causes the broad emission linewidth observed in the SQW investigated here. Additionally, interface roughness, dislocations caused by lattice and thermal coefficient mismatches between InGaIn and AlGaIn, and native defects may also contribute to the emission linewidth broadening, which further leads to the difficulties in resolving different emission lines in cw PL spectra.

The excitation intensity dependencies of the recombination lifetimes τ_1 and τ_2 have also been measured and are shown in Fig. 8. Both τ_1 and τ_2 are almost constants when I_{exc} is varied by two orders of magnitude. This suggests that the localized exciton recombination is always the dominant

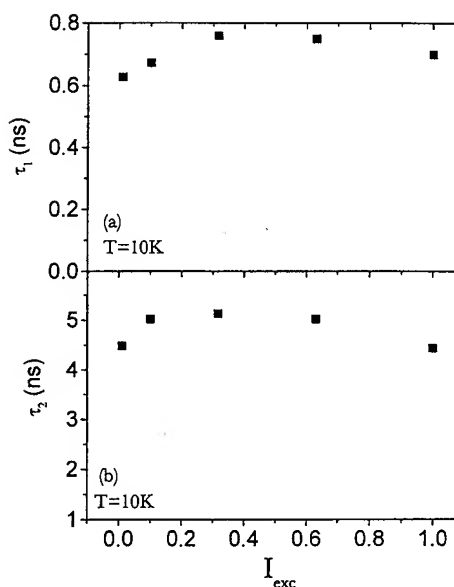


FIG. 8. Excitation intensity dependencies of the recombination lifetimes τ_1 and τ_2 measured at the spectral peak positions at 10 K.

optical transition in InGaIn/AlGaIn SQW samples studied here.

V. SUMMARY

Mechanisms of optical transitions of InGaIn/AlGaIn SQW have been studied by time-resolved PL under different conditions. The dynamical behavior of the PL emission reveals that the main emission line in these SQW is due to the combination of the localized exciton and a band-to-impurity recombination. The spectral line shape and the recombination dynamics of the localized exciton and the band-to-impurity transitions have been systematically investigated at different conditions.

ACKNOWLEDGMENTS

The research at Kansas State University is supported by ARO (DAAH04-96-1-0371), BMDO/ONR (N00014-96-1-0885), and NSF (DMR-95-28226). The research at North Carolina State University is supported by ARO (DAAH04-96-1-0173) and ONR (N00014-92-J-1473).

¹H. Morkoc, S. Strite, G. B. Gao, M. E. Lin, B. Sverdlov, and M. Burns, *J. Appl. Phys.* **76**, 1363 (1994).

²S. Nakamura, M. Senoh, N. Iwasa, and S. Nagahama, *Jpn. J. Appl. Phys.* **34**, L797 (1995).

³S. Nakamura, M. Senoh, N. Iwasa, S. Nagahama, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto, *Appl. Phys. Lett.* **68**, 2105 (1996).

⁴M. Smith, J. Y. Lin, H. X. Jiang, A. Salvador, W. Kim, A. Botchkarev, and H. Morkoc, *Appl. Phys. Lett.* **69**, 2453 (1996).

⁵M. Smith, G. D. Chen, J. Y. Lin, H. X. Jiang, M. A. Khan, and Q. Chen, *Appl. Phys. Lett.* **69**, 2837 (1996).

⁶C. K. Sun, S. Keller, G. Wang, M. S. Minsky, J. E. Bowers, and S. P. DenBaars, *Appl. Phys. Lett.* **69**, 1936 (1996).

⁷E. S. Jeon, V. Kozlov, Y. K. Song, A. Vertikov, M. Kuball, A. V. Nur-mikko, H. Liu, C. Chen, R. S. Kern, C. P. Kuo, and M. G. Craford, *Appl. Phys. Lett.* **69**, 4194 (1996).

⁸W. Li, P. Bergman, B. Monemar, H. Amano, and I. Akasaki, *J. Appl. Phys.* **81**, 1005 (1997).

⁹G. D. Chen, M. Smith, J. Y. Lin, H. X. Jiang, A. Salvador, B. N. Sverdlov, A. Botchkarev, and H. Morkoc, *J. Appl. Phys.* **79**, 2675 (1996).

¹⁰P. Avouris and T. N. Morgan, *J. Chem. Phys.* **74**, 4347 (1981).

¹¹M. Smith, G. D. Chen, J. Y. Lin, H. X. Jiang, A. Salvador, B. N. Sverdlov, A. Botchkarev, H. Morkoc, and B. Goldenberg, *Appl. Phys. Lett.* **68**, 1883 (1996).

Theory of interfaces and surfaces in wide-gap nitrides

Marco Buongiorno Nardelli,^{a)} Krzysztof Rapcewicz, and J. Bernholc

Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 29 March 1997; accepted 22 April 1997)

A selection of the results of a theoretical investigation of the properties of interfaces and surfaces of the wide-gap III-V nitride semiconductors is reviewed. The electronic properties of wurtzite heteroepitaxial interfaces of AlN and GaN, incorporating the effects of strain, are discussed. In particular, we find that this interface is of type I and have calculated the valence-band offset to be -0.57 eV. The surface energies and atomic geometries of the 2×2 reconstructions of the (0001) face of GaN are also presented. In conditions which are rich in a given species, an adatom reconstruction of that species is found to be the most energetically favorable: for gallium-rich conditions, the reconstruction with a gallium adatom on a T_3 site is the most stable, while for nitrogen-rich conditions the reconstruction with a nitrogen adatom on the H_3 site is energetically the most favorable. © 1997 American Vacuum Society. [S0734-211X(97)11904-7]

I. INTRODUCTION

Semiconductor heterojunctions find wide applications in microelectronics. These include lasers, photodetectors, high-efficiency solar cells, and radiation-resistant integrated circuits.¹ The III-V nitrides with their wide band gaps, high thermal stability and conductivity, and radiation resistance have long been viewed as attractive candidates for use in microelectronic devices.²

The recent demonstration of a GaN-based blue laser, which follows three decades of research, has served to highlight the potential of these materials.³ However, at present, the mechanisms underlying the workings of the blue laser are not understood. To this end and in order to develop other nitride-based devices, great effort is being expended to understand and control the materials properties of the wide gap nitrides.

II. METHODOLOGY

The standard *ab initio* plane-wave pseudopotential method⁴⁻⁶ was used in the calculations of interface properties. To ensure convergence of the nitrogen pseudopotential, an energy cut-off for the plane-wave expansion of 50 Ry was chosen. The equivalent of 10 k -points for bulk and superlattice calculations in the zinc-blende structure⁷ and 6 k -points for calculations of the wurtzite structure⁸ were employed. A careful check of the convergence in both the size of the plane-wave basis and the number of special points was made. For the surface calculations, a multi-grid-based total-energy method that uses a real-space grid as the basis was used. This approach has been used to study a wide range of systems involving large numbers of atoms.⁹ The ions were relaxed using the "fast-relax" algorithm.¹⁰

The Perdew-Zunger parametrization¹¹ of the Ceperley-Alder form¹² of the exchange-correlation energy was chosen. Nonlocal, norm-conserving pseudopotentials¹³⁻¹⁵ were included using the Kleinman-Bylander approach.¹⁶ For gallium, a recently developed norm-conserving pseudopotential,

which includes a nonlocal core correction,¹⁷ and permits an efficient description without the need for an explicit treatment of the d valence electrons was used. For nitrogen, a standard pseudopotential with a neutral configuration as the atomic reference for all states was employed. These pseudopotentials have been demonstrated to reproduce accurately the bulk properties of GaN.¹⁸

The calculated bulk properties are presented in Table I. The theoretical lattice parameters of both the zinc-blende and wurtzite forms agree very well with the experiment; a similar level of accuracy is expected for the surface and interface calculations described below. The cohesive energy of GaN and α -Ga is 10.42 eV and 3.4 eV, respectively. The binding energy of nitrogen molecule is 5.87 eV per nitrogen atom, in good agreement with other local density approximation (LDA) calculations.¹⁹ The theoretical heat of formation of GaN is thus 1.15 eV, in good agreement with the experimental value of 1.14 eV.

III. INTERFACE PROPERTIES

Interface behavior is very important in determining the properties of semiconductor devices. We have investigated the wurtzite (0001) GaN/AlN interface. Superlattices of AlN and GaN are expected to be strained because of the 3.5% theoretical lattice mismatch between AlN and GaN (the experimental value is 2.7%). The effects of the strain were incorporated using macroscopic elasticity theory.²⁰ In this theory, each half of the heterojunction is treated as a strained bulk with a fixed in-plane lattice constant $a_{||}$. The strain energy of the system is minimized, keeping the in-plane lattice constant fixed, to determine the perpendicular lattice constant, c , of the epilayer. For AlN and GaN, the calculated elastic constants used to determine the strains are given in Table II. Total energy calculations confirm that macroscopic elasticity theory predicts well the perpendicular lattice constant for the epilayer. The residual relaxation of the atoms at the interface is negligible (<0.05 Å) and does not affect the band offset of the interface.¹⁸

^{a)}Electronic mail: nardelli@nemo.physics.ncsu.edu

TABLE I. Calculated bulk properties of zinc-blende and wurtzite nitride semiconductors. The values of the gap at the Γ -pt (E_g) and of the valence-band width (ΔE_{vbw}) are the LDA results. Note that the LDA indirect gap in zinc-blende AlN is 3.2 eV. Experimental values are in brackets and follow Ref. 35.

	AlN	GaN	InN
zinc-blende			
a_0 (Å)	4.37 (4.38)	4.52 (4.5)	5.01 (4.98)
B_0 (Mbar)	2.02 (2.02)	1.70 (1.90)	1.58 (1.37)
wurtzite			
a (Å)	3.09 (3.11)	3.20 (3.19)	3.55 (3.54)
c/a	1.62 (1.60)	1.63 (1.63)	1.63 (1.61)
u (units of c)	0.378 (0.382)	0.376 (0.377)	0.375
B_0 (Mbar)	1.99 (2.02)	1.69 (1.95, 2.37)	1.62 (1.26, 1.39)

The band offsets of the (0001) GaN/AlN strained heterojunction were studied following the procedure of Ref. 21. The valence band offset is divided into a band-structure contribution, which is the difference between the energies of the valence-band and conduction-band edges when the average electrostatic potentials of the epilayers are aligned, and the difference in the value of the average electrostatic potential (ΔV) in the two epilayers of the heterostructure. Strain will affect the electronic properties of the interface through the variation of the average electrostatic potentials and through the variations of the band edges (deformation potentials). The calculated valence-band offset in the case of an AlN in-plane lattice constant (strained GaN) is estimated to be -0.57 eV,²² a value smaller than the result for the strained nonpolar (001) GaN/AlN interface. The ratio of the conduction-band to valence-band offset is 65:20. These results agree very well with the experimental measurements of the (0001) wurtzite interface.²³

As spin-orbit effects in both AlN and GaN have been shown to be of the order of 20 meV, their difference, which enters into the calculations of the band offsets, is much smaller than other possible sources of systematic error.²⁴ These include the neglect of the anion p - and cation d -state repulsion²⁵ and the well-known neglect of many-body effects in the LDA. In GaN/AlN interfaces, the inclusion of the $3d$ electrons as valence electrons results in a constant shift of 0.2 eV, which is less than the experimental error,²³ and does not change the character of the interface. Incorporating this shift gives results in agreement with previous work using a d valence pseudopotential³⁰ and an all-electron calculation.²⁶ The importance of many-body effects on the band offsets is not known and awaits a future GW calculation.

TABLE II. Elastic constants calculated for zinc-blende AlN, GaN, and InN (in units of Mbar). The experimental values in brackets are from Ref. 36.

	AlN	GaN	InN
c_{11}	3.01	2.61 (2.64)	2.14
c_{12}	1.62	1.27 (1.53)	1.37

The wurtzite system displays pyroelectric and piezoelectric behavior.²⁷ These effects will manifest themselves macroscopically in multiple quantum wells along those directions that do not have a perpendicular mirror plane.^{28,29} In (0001) strained GaN/AlN, we observe a substantial electric field as has been previously noted by Satta and co-workers.³⁰ We have calculated the spontaneous bulk polarization of unstrained AlN and the strain-induced polarization for the GaN epilayer, in order to distinguish the bulk pyroelectric and piezoelectric contributions to this field from that induced by the interface. A superlattice consisting of eight layers of wurtzite and six layers of lattice-matched zinc-blende was used.³¹ In unstrained zinc-blende, the bulk spontaneous polarization is rigorously zero. Further, zinc-blende-wurtzite interface does not introduce any chemical or geometrical perturbations. This construction, therefore, permits the unambiguous determination of the spontaneous polarization present in the wurtzite structure from the slope of the macroscopic average of the electrostatic potential. The spontaneous polarization (P_3) of AlN and GaN in equilibrium is $-1.227 \mu\text{C}/\text{cm}^2$ and $-0.448 \mu\text{C}/\text{cm}^2$, respectively; the polarization of the strained GaN is $-0.454 \mu\text{C}/\text{cm}^2$. These values are comparable to the computed bulk polarization in BeO.³¹ The effect of the interface dipole is small; the polarization in the GaN/AlN multiple quantum well is well described as the sum of the polarizations of the constituent epilayers of the multiple quantum well. The estimated contribution of the interface dipole (which includes the response of one epilayer to the field of the other) is $0.057 \mu\text{C}/\text{cm}^2$, which is of opposite sign and an order of magnitude smaller than the bulk polarizations. The computed value of the polarization in the superlattice agrees with that estimated from experiment by Martin *et al.*²³

IV. SURFACE RECONSTRUCTIONS

The (0001) surface is a polar surface. The ideal surface has dangling bonds with charge distributions that are energetically unfavorable. If kinetically permitted, the surface atoms will relax and redistribute the "dangling-bond" charge density so as to satisfy the valences of all of the surface species. This can be achieved via structural relaxation (with a concomitant charge transfer between surface atoms). Recent experimental studies have revealed the presence of 2×2 reconstructions during growth.³² In this article, we restrict our attention to these reconstructions on the gallium-terminated face.

We considered eight different 2×2 reconstructions, namely the ideal relaxed structure and gallium vacancy, nitrogen adatom, gallium adatom, and nitrogen trimer, each on the T_3 and H_3 (hollow) sites (see Figs. 1 and 2). We find that adatom reconstructions are the most energetically favorable of the studied reconstructions. In particular, in the gallium-rich case that the gallium adatom on the T_3 site has the lowest energy, while in the nitrogen-rich case the nitrogen adatom on the H_3 is the most stable.

The slabs contained four bilayers of GaN, three bilayers of which were relaxed. Relaxations in the third bilayer were

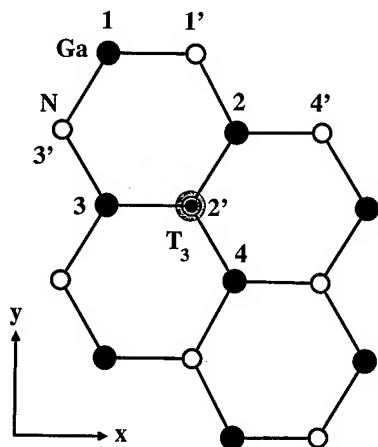


FIG. 1. Schematic top view of the 2×2 gallium-atom reconstruction on the (0001) surface of GaN. The gallium adatom (in grey) sits above a nitrogen in the second layer on the T_3 site. The numbers denote gallium atoms on the surface, while the primed numbers denote nitrogen atoms in the first subsurface layer.

negligible, indicating that a sufficiently thick slab was employed. Pseudo-hydrogens with $Z=0.75$ passivated the nitrogen-terminated face.³³ If the work functions of the two surfaces are different, then the periodic boundary conditions of the supercell enforce a common electrostatic potential in the vacuum and result in an unphysical change in the electric potential in the vacuum region equal to the difference between the two work functions. A large vacuum region of 10 \AA served to reduce the size of this field. Previous calculations that explicitly included a correction for this field did not show significant differences in the final equilibrium geometries,³⁴ and consequently field corrections were not included in our calculations. An orthorhombic supercell with grid spacing of 0.21, 0.22, and 0.20 a.u. in the x , y , and z directions, respectively, was employed. The eight reconstructions mentioned above were studied using Γ -point sampling. For the lowest energy structures of each class, i.e., nitrogen

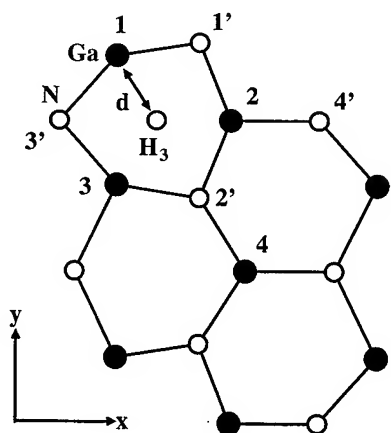


FIG. 2. Schematic top view of the 2×2 nitrogen-atom reconstruction on the (0001) surface of GaN. The nitrogen adatom sits above the hollow (H_3) site. The numbers denote gallium atoms on the surface, while the primed numbers denote nitrogen atoms in the first subsurface layer.

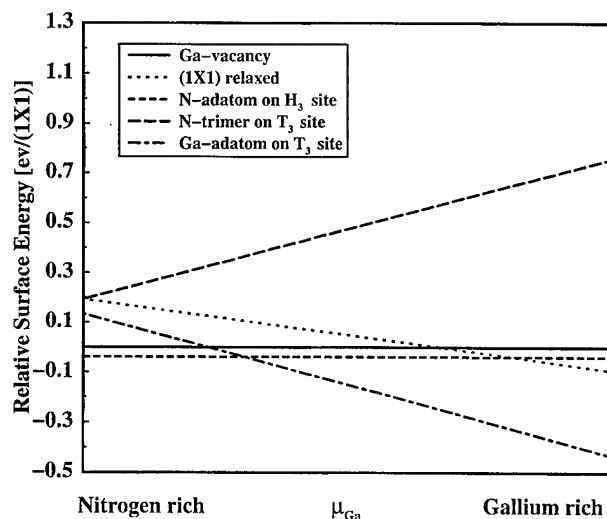


FIG. 3. Formation energy vs gallium chemical potential for the (0001) surface. The maximum chemical potential for N (Ga) is equal to the energy per atom calculated for N_2 (bulk Ga). Two k -points are sampled in the irreducible Brillouin zone.

adatom on H_3 site, nitrogen trimer on the T_3 site, etc., a further calculation using two k -points in the irreducible Brillouin zone was carried out.

Figure 3 displays the relative surface energies for the five different reconstructions calculated using 2 k -points in the irreducible Brillouin zone. Under gallium-rich conditions and for most of the range of the chemical potential, the lowest energy reconstruction is the gallium adatom on the T_3 site, while under nitrogen-rich conditions, the reconstruction with the nitrogen adatom on the H_3 site is preferred energetically.

The gallium-adatom reconstruction is shown in Fig. 1. The adatom sits on the T_3 site above the subsurface nitrogen atom; the gallium-nitrogen distance is 2.46 \AA . The gallium-adatom surface-gallium bond length is 2.40 \AA , which is very close to the Ga dimer distance in α -gallium (2.44 \AA). The bond angle between two gallium surface atoms and the adatom is 82° . The in-plane relaxation of the surface atoms is negligible; the only observed relaxation of the first bilayer is along the z direction. The three gallium atoms bonded to the adatom remain coplanar to a significant degree, while the remaining unbonded gallium surface atom relaxes into the slab. The proximity of the gallium adatom to the nitrogen subsurface atom lowers the electrostatic energy and makes this particular configuration energetically more favorable than the gallium adatom reconstruction with the adatom on the hollow site.

The nitrogen-adatom reconstruction is shown in Fig. 2. The adatom sits on the hollow (H_3) site. The nitrogen-adatom surface-gallium bond length is 2.0 \AA and is very close to the Ga-N bond length distance in GaN bulk (1.96 \AA). The bond angle between two gallium surface atoms and the nitrogen adatom is 89° . The relaxation perpendicular to the surface is very similar to that in the gallium-adatom reconstruction. This configuration is stabilized because it separates the nitrogen adatom from the subsurface nitrogen atom and thereby lowers the electrostatic repulsion.

V. SUMMARY

We have reviewed a selection of theoretical results for the properties of interfaces and surfaces of the gallium and aluminium nitrides. The electronic properties of wurtzite heteroepitaxial interfaces of AlN and GaN incorporating the effects of strain were discussed. The strained AlN/GaN interface is of type I and has a valence-band offset of -0.57 eV. As a result of the pyro- and piezoelectric nature of the wurtzite nitrides, macroscopic polarization of $0.057 \mu\text{C}/\text{cm}^2$, is present in multiple quantum wells along the (0001) direction. The surface energies and atomic geometries of the 2×2 reconstructions of the (0001) face of GaN were also investigated. Adatom reconstructions were found to be the most energetically favorable over the whole range of gallium chemical potential: under gallium-rich conditions, the reconstruction with a gallium adatom on a T_3 site is the most energetically favorable, while under nitrogen-rich conditions the reconstruction with a nitrogen adatom on the H_3 site has the lowest energy.

¹R. Bauer and G. Margaritondo, *Phys. Today* **40**, 27 (1987).

²M.-E. Lin, B. N. Sverdlov, S. Strite, H. Morkoç, and A. E. Drakin, *Electron. Lett.* **29**, 1759 (1993).

³S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto, *Jpn. J. Appl. Phys.* **1** **35**, L74 (1996).

⁴M. Payne, M. Teter, D. Allan, T. Arias, and J. Joannopoulos, *Rev. Mod. Phys.* **64**, 1045 (1992).

⁵W. Pickett, *Comput. Phys. Rep.* **9**, 115 (1989).

⁶D. Remler and P. Madden, *Mol. Phys.* **70**, 921 (1990).

⁷S. Froyen, *Phys. Rev. B* **39**, 3168 (1989).

⁸D. J. Chadi and M. L. Cohen, *Phys. Rev. B* **8**, 5747 (1973).

⁹E. L. Briggs, D. J. Sullivan, and J. Bernholc, *Phys. Rev. B* **52**, R5471 (1995); E. L. Briggs, D. L. Sullivan, and J. Bernholc, *ibid.* **54**, 14362 (1996).

¹⁰C. Wang, Q.-M. Zhang, and J. Bernholc, *Phys. Rev. Lett.* **69**, 3789 (1992).

¹¹J. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).

¹²D. Ceperley and B. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).

¹³G. Bachelet, D. Hamann, and M. Schlüter, *Phys. Rev. B* **26**, 4199 (1982).

¹⁴D. Hamann, M. Schlüter, and C. Chiang, *Phys. Rev. Lett.* **43**, 1494 (1979).

¹⁵D. Hamann, *Phys. Rev. Lett.* **40**, 2980 (1989).

¹⁶L. Kleinman and D. Bylander, *Phys. Rev. Lett.* **48**, 1425 (1982).

¹⁷S. Louie, S. Froyen, and M. L. Cohen, *Phys. Rev. B* **26**, 1738 (1982).

¹⁸M. Buongiorno Nardelli, K. Rapcewicz, and J. Bernholc, *Phys. Rev. B* **55**, R7323 (1997).

¹⁹B. G. Johnson, P. Gill, and J. A. Pople, *J. Chem. Phys.* **98**, 5613 (1993).

²⁰C. Van de Walle and R. M. Martin, *Phys. Rev. B* **34**, 5621 (1986).

²¹A. Baldereschi, S. Baroni, and R. Resta, *Phys. Rev. Lett.* **61**, 734 (1988).

²²The band-offset was calculated following A. Munoz, N. Chetty, and R. M. Martin, *Phys. Rev. B* **41**, 2976 (1990). The position of the interface was unambiguously determined from the position of the interface dipole.

²³G. Martin, A. Botchkarev, A. Rockett, and H. Morkoç, *Appl. Phys. Lett.* **65**, 610 (1994); **68**, 2541 (1996).

²⁴M. Suzuki, T. Uenoyama, and A. Yanase, *Phys. Rev. B* **52**, 8132 (1995).

²⁵S. Wei and A. Zunger, *Phys. Rev. Lett.* **59**, 144 (1987).

²⁶E. Albanesi, W. Lambrecht, and B. Segall, *J. Vac. Sci. Technol. B* **12**, 2470 (1994).

²⁷N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, PA, 1976), Ch. 27.

²⁸D. Smith, *Solid State Commun.* **57**, 919 (1986).

²⁹A. Bykhovski, B. Gelmont, and M. Shur, *Appl. Phys. Lett.* **63**, 2243 (1993); *J. Appl. Phys.* **74**, 6734 (1993).

³⁰A. Satta, V. Fiorentini, A. Bosin, F. Meloni, and D. Vanderbilt, *Mater. Res. Soc. Symp. Proc.* **395**, 515 (1996).

³¹M. Posternak, A. Baldereschi, A. Catellani, and R. Resta, *Phys. Rev. Lett.* **64**, 1777 (1990).

³²M. A. L. Johnson, Z. Yu, C. Boney, W. H. Rowland, Jr., W. C. Hughes, J. W. Cook, Jr., J. F. Schetzina, N. A. El-Masry, M. T. Leonard, H. S. Kong, and J. A. Edmond, *Mater. Res. Soc. Symp. Proc.* **449**, 271 (1997).

³³K. Shiraiishi, *J. Phys. Soc. Jpn.* **59**, 3455 (1990).

³⁴J. Neugebauer and M. Scheffler, *Phys. Rev. B* **46**, 16067 (1992).

³⁵A. Wright and J. Nelson, *Phys. Rev. B* **50**, 2159 (1994); **51**, 7866 (1995).

³⁶K. Kim, W. Lambrecht, and B. Segall, *Phys. Rev. B* **50**, 1502 (1994).

Growth and characterization of light emitting ZnS/GaN heterostructures

E. C. Piquette, Z. Z. Bandić, J. O. McCaldin, and T. C. McGill^{a)}
Thomas J. Watson, Sr. Laboratory of Applied Physics, California Institute of Technology,
Pasadena, California 91125

(Received 13 January 1997; accepted 10 April 1997)

Heterostructures involving ZnS/GaN show promise for the injection of holes from *p*-GaN into *n*-ZnS. Utilizing knowledge obtained from ZnS phosphor technology, this combination could result in a new type of multi-color electroluminescent display. Further, this combination provides a very interesting interface. Both ZnS and GaN are very ionic materials. Hence, it is desirable that the interface will be relatively benign, and that charge injection can occur despite the large lattice mismatch and resulting misfit defects that form near the interface. The highly lattice mismatched structures ZnS/GaN and ZnS/Al₂O₃ were grown by molecular beam epitaxy using elemental sources. Growth rates of up to 0.4 μm ph were observed for the lower growth temperatures, with rapidly diminishing rates for temperatures above 350 °C. The GaN substrate consisted of a 3 μm epilayer grown on sapphire (0001) by metalorganic chemical vapor deposition. Reflection high energy electron diffraction observations indicate that the zincblende ZnS layers commonly contain (111) twins, although films with no visible twin spots could be grown at a high substrate temperature. The sulfide layers were characterized using photoluminescence, x-ray diffraction, and scanning electron microscopy (SEM). X-ray peaks typically had widths of 400 arcsec for $\omega/2\theta$ scans, and somewhat worse for ω scans, indicating mosaic tilt. Photoluminescence spectra of the ZnS films doped with Ag and Al demonstrated the characteristic blue transition near 440 nm. SEM electron channeling patterns indicated that the ZnS films lie at a 30° rotated orientation to the Al₂O₃ (0001) substrate, as does GaN. *p*-GaN/*n*-ZnS devices were fabricated using standard photolithography techniques. The current voltage characteristics are reported and preliminary electroluminescence results are discussed for this heterojunction system. © 1997 American Vacuum Society. [S0734-211X(97)06804-2]

I. INTRODUCTION

Zinc sulfide, with its wide direct band gap, has long been known as a versatile and efficient light emitting compound. It has been extensively used as a phosphor, and continues to show promise in semiconductor light emitting devices, such as ZnS based flat panel displays,¹⁻³ light emitting diodes,⁴ and semiconductor lasers.⁵ However, difficulties in producing high quality doped material, particularly *p*-type, have otherwise reduced the utility of the sulfide system.

Recently, high conductivity *p*-type GaN has become available.⁶⁻⁸ With a valence band edge below that of ZnS,⁹ it has been proposed that GaN may be a good hole injector into ZnS and other light emitting semiconductor materials.¹⁰ To this end, we have grown ZnS on GaN (0001) and sapphire (0001) substrates by molecular beam epitaxy (MBE) in an attempt to characterize this potentially interesting system as well as gain insight into highly lattice mismatched systems in general.

While there have been many reports of ZnS grown by MBE on near lattice matched substrates such as Si,^{11,12} GaP,¹³⁻¹⁵ and GaAs,¹⁶⁻¹⁸ with lattice mismatches of 0.4%, 0.8%, and 4%, respectively, little has been reported of epitaxial growth on the highly mismatched substrates GaN ($\sim +20\%$ mismatch) or Al₂O₃ ($\sim -20\%$ mismatch). These interfaces are very interesting because of the large mismatch

and the unusual properties of the constituents. First, it is not expected that the ZnS epilayer will grow coherently strained for more than 1 monolayer. More likely, relaxation in the layers closest to the interface will result in dangling bonds and numerous dislocations. In fact tilting and three dimensional (3D) relaxation are likely the most energetically preferred strain relaxation mechanisms.¹⁹ Second, GaN and ZnS are known to be good light emitters even when the defect density is high. Hence, this interface provides us with a potentially technologically useful junction even though it is likely to be heavily defected.

II. MBE GROWTH AND RHEED OBSERVATIONS

ZnS epilayers were grown in a modified Perkin-Elmer 430 MBE system, equipped with a valved cracking source for sulfur (6 N purity), Knudsen cell sources for elemental Zn (6 N), Al (6 N), Ag (6 N), and a radio frequency nitrogen (5 N5) plasma source (Oxford Applied Research). Except where noted, the sulfur cracking zone temperature was held at 700 °C. The growth chamber maintains a base pressure during growth of 3×10^{-10} Torr with cryopaneling fully cooled. Samples were either bonded with In to molybdenum holder blocks, or secured to similar blocks with Mo or Si clips. We found that In bonding provided better temperature uniformity and reproducibility than the use of clips. Temperatures were measured by thermocouple, and reflection high energy electron diffraction (RHEED) operating at

^{a)}Electronic mail: tcm@ssdp.caltech.edu

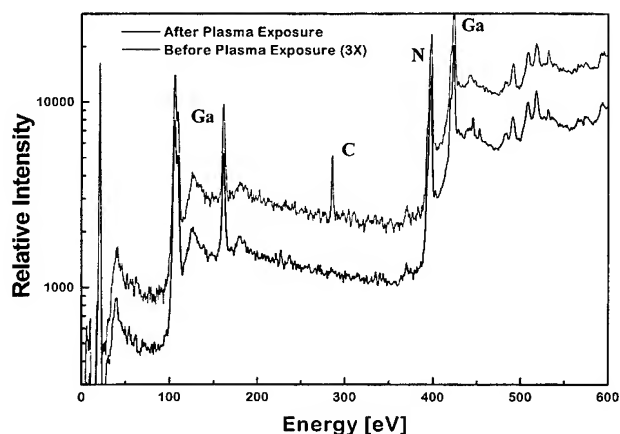


FIG. 1. X-ray photoemission spectrum of GaN before and after exposure to *in situ* rf nitrogen plasma.

10 keV was used as a characterization tool during growth. Some ZnS films were doped with Ag and Al.

The GaN substrates consisted of a 3- μm -thick epilayer grown by metalorganic chemical vapor deposition (MOCVD) on *c*-plane sapphire. Van der Pauw Hall measurements performed on the GaN layers indicated *p*-type conductivity of 5 $\Omega\text{ cm}$, a carrier concentration of $\sim 2 \times 10^{17}\text{ cm}^{-3}$, and a mobility of 7 $\text{cm}^2/(\text{V s})$. Sapphire (0001) substrates were sourced from Union Carbide Crystal Products. Substrates were degreased in organic solvents and loaded into the MBE vacuum system. Before epilayer growth, the samples were transferred *in situ* between the growth chamber and an x-ray photoemission spectroscopy (XPS) analysis chamber. XPS analysis on the GaN layers revealed carbon surface contamination of atmospheric origin. It was found that the carbon could be removed by exposing the sample to a rf nitrogen soak at a substrate temperature of 500 $^{\circ}\text{C}$ for ~ 20 min (Fig. 1). The plasma operated at a chamber pressure of 1×10^{-6} Torr and 200 W rf power. This treatment also visibly improved the RHEED pattern of the GaN. Contamination could be reduced, but not eliminated, by thermal cleaning alone. The sapphire substrates were thermally treated at 500 $^{\circ}\text{C}$ for 30 min prior to growth, after which sharp RHEED streaks were manifest. We found that etching of substrates made no noticeable difference.

Preceding MBE growth the GaN and sapphire substrates showed sharp RHEED streaks. After opening the source shutters the RHEED pattern quickly became spotty for most films, suggesting that growth proceeds in the 3D (Volmer-Weber) mode. The spotty pattern persisted throughout the growth period, and indicates that the ZnS films are of the zincblende structure with (111) orientation and are single crystalline or composed of mosaic crystallites which are aligned with the underlying substrate lattice. Most films displayed twinning in the (111) growth plane [Fig. 2(B)] which is commonly observed for growth on a (111) oriented substrate,^{13,18} although this twinning was absent and replaced by (11 $\bar{1}$) twins for some films grown with a lower sulfur cracking temperature of 300 $^{\circ}\text{C}$ [Fig. 2(C)]. Films grown at

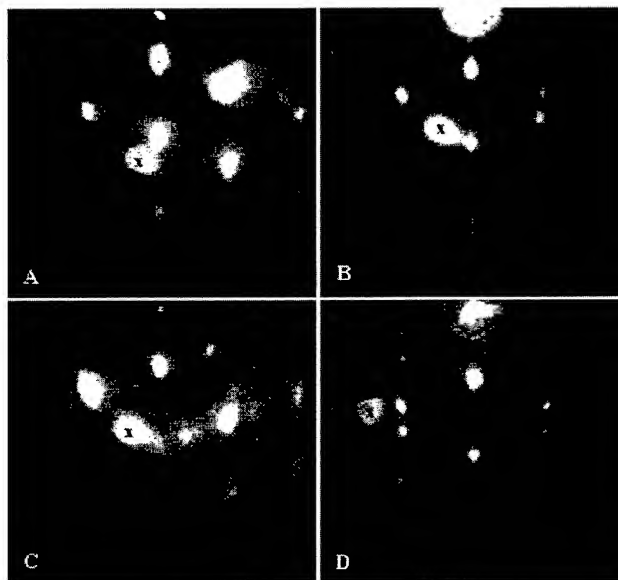


FIG. 2. 10 keV RHEED images [(1 $\bar{1}0$) azimuth] of ZnS on sapphire (A) grown at 400 $^{\circ}\text{C}$ with no visible twin spots, (B) grown at 250 $^{\circ}\text{C}$ with sulfur cracking zone temperature at 700 $^{\circ}\text{C}$ and showing (111) twins, (C) grown at 250 $^{\circ}\text{C}$ with cracking zone at 300 $^{\circ}\text{C}$ showing (11 $\bar{1}$) twins, and (D) ZnS:Al,Ag grown on GaN [(111) twinned]. The smudge common to all pictures is due to a defect in the phosphor screen (x).

the higher temperature of 400 $^{\circ}\text{C}$ showed no visible twinning [Fig. 2(A)]²⁰ and show superior surface morphology, although they suffer from extremely low growth rates. Anti-thetically, ringlike patterns were observed for films grown at very low temperatures or with poor surface preparation. Streaky patterns were not observed for any of the samples grown at temperatures between 150 $^{\circ}\text{C}$ and 400 $^{\circ}\text{C}$ and with source flux beam equivalent pressures (BEPs) of $(0.3-2.0) \times 10^{-7}$ Torr.

III. FILM CHARACTERIZATION

Film thicknesses were measured by depth profiler and by spectroscopic ellipsometry, and growth rates were calculated assuming a constant rate throughout the growth period. Figure 3 shows the calculated rates as a function of growth temperature. Rates of up to 0.4 $\mu\text{m/h}$ were observed for the lower growth temperatures, with rapidly diminishing rates for temperatures above 350 $^{\circ}\text{C}$. This trend is in agreement with other work.^{11,13,18} A typical film thickness was 600 nm.

Some films were characterized by scanning electron microscopy (SEM) operating in electron channeling pattern (ECP) mode. The ECP scans are useful as an indication of crystal symmetry, quality, and orientation. Channeling patterns were visible for as grown films grown at 400 $^{\circ}\text{C}$ and for post-annealed films. The ECP shown in Fig. 4 is of a sample annealed in sulfur over pressure at 1000 $^{\circ}\text{C}$.²¹ For films of poorer crystallinity as quantified by x-ray measurements, no ECP was visible. As confirmed by RHEED observations, the scan shows that the zincblende ZnS grows at a 30 $^{\circ}$ rotated orientation to the sapphire substrate lattice. That is, the (111) faces of ZnS parallel the (0001) faces of Al_2O_3 and the [1 $\bar{1}0$]

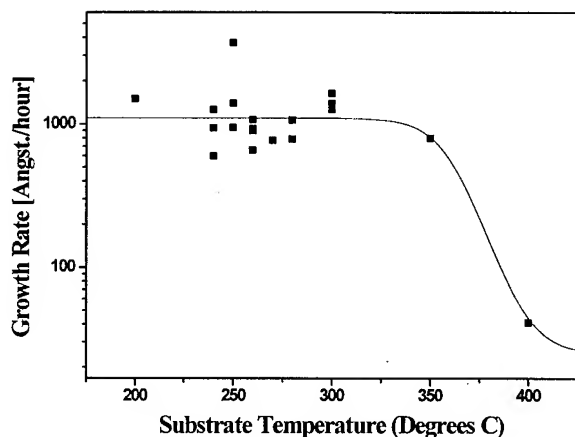


FIG. 3. Temperature dependence of the growth rate of ZnS on sapphire (0001). The line is meant to guide the eye.

azimuth points along the $[10\bar{1}0]$ direction (Miller-Bravais notation) of the Al_2O_3 . GaN has also been observed to grow in a 30° oriented state on sapphire (0001). The degree of detail visible in the ECP also indicates good crystal quality.

High resolution x-ray diffraction analysis was performed using a four crystal $[\text{Ge} (220)]$ diffractometer and $\text{Cu } K\alpha$ x rays. Results indicated that the films are composed of mosaic crystalline domains which are slightly tilted and rotated from one another. Such mosaic films are commonly observed in other materials systems where large lattice mismatch exists, such as Si on Al_2O_3 , and GaAs on Si. The full width at half maximum (FWHM) of (111) peaks were consistently close to 400 arcsec for $\omega/2\theta$ scans, invariable with growth conditions. X-ray rocking curves (ω scans) showed peak widths of up to several degrees, depending on growth temperature, film thickness, and doping (Fig. 5). The FWHM of ω scans provides an indication of the degree of mosaic tilting that is present in the film. Undoped samples grown at 400°C

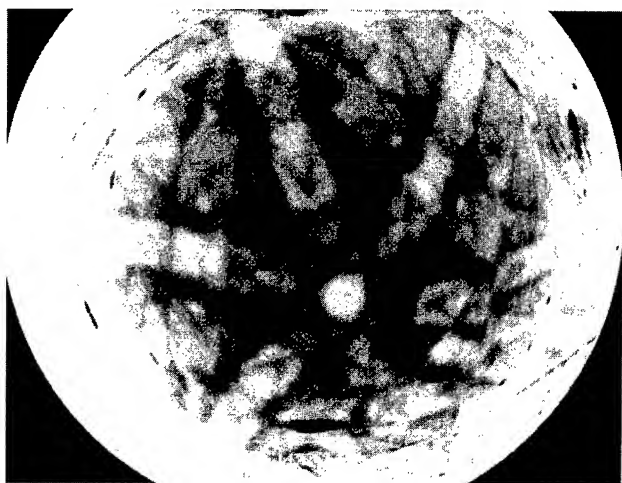


FIG. 4. Shown is the SEM electron channeling pattern of ZnS grown on sapphire (0001), post-annealed at 1000°C under high sulfur over pressure. The ECP indicates that the ZnS layer lies rotated to the substrate. Distortion in the figure is caused by charging effects.

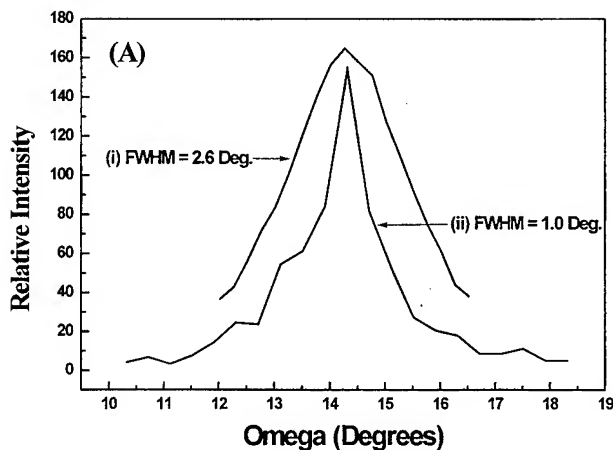


FIG. 5. X-ray rocking curves for ZnS grown on sapphire at (i) 250°C and (ii) 400°C .

showed the best as grown rocking curve FWHM of ~ 60 arcmin. Further improvement in crystalline quality was obtained by post-annealing at 1000°C in a high pressure sulfur environment.²¹

Low temperature (5 K) photoluminescence (PL) measurements were made using the 325 nm wavelength of a HeCd laser. The samples doped with Ag and Al produced a bright blue luminescence with spectra as shown in Fig. 6. Visible are the characteristic silver high (390 nm) and low (440 nm) energy emission bands.²² The relative intensities of the two bands are known to be dependent on sample preparation. In the samples studied, both bands are present for the ZnS:Al,Ag films grown on sapphire, while the lower energy peak is dominant for the films grown on n -GaN, and the film grown on p -GaN shows only emission near the Ag higher energy wavelength. The p -GaN layer also has luminescence in the blue-violet region, and is not resolvable from the ZnS emission shown in Fig. 6(A). This variation in Ag transition peak intensities is not well understood at this time, and further study is warranted. Also present in Fig. 6(B) is the near band edge excitonic structure (350–400 nm) of the unintention-

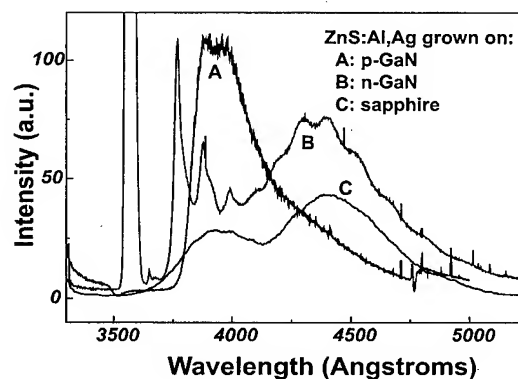


FIG. 6. Low temperature (5 K) photoluminescence spectra of ZnS doped with silver and aluminum on (A) p -GaN, (B) unintentionally doped n -GaN, and (C) sapphire (0001).

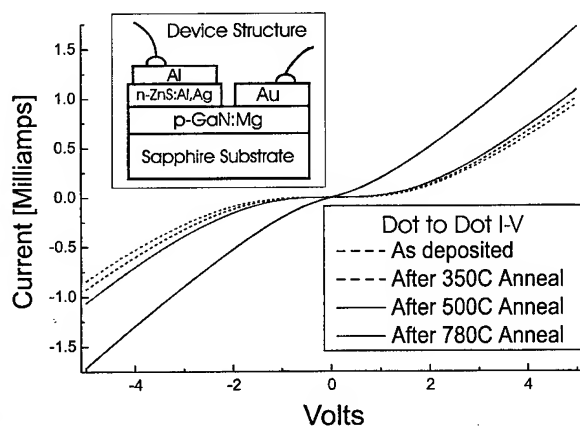


FIG. 7. Dot to dot current voltage characteristic of p -GaN as sputter deposited, and after subsequent anneals at 350 °C, 500 °C, and 780 °C for 1 min in nitrogen gas. The inset diagram shows the schematic structure of the fabricated p -GaN/ n -ZnS device.

tionally doped n -type GaN substrate. The ZnS band gap is too wide to probe near band edge features with this laser.

IV. DEVICE FABRICATION

p -GaN/ n -ZnS devices were fabricated using standard photolithographic processes. The insulating sapphire substrate necessitated the use of lateral contacts. The fully processed structure is shown in the inset of Fig. 7. An aluminum top contact was first sputter deposited on the as grown ZnS:Al,Ag/GaN:Mg layers. Mesas were then patterned and the Al and ZnS layers were etched in a solution of Br:HBr:H₂O (1:100:100), which was found not to quickly etch the GaN. Au p -type contacts were then patterned and sputter deposited using a lift-off process. It is not believed that either contact is ohmic. The effects of annealing on the dot-to-dot current-voltage curves for p -GaN/Au contacts were observed on a separate substrate, and the results are presented in Fig. 7. As deposited, it is seen that the contacts are highly rectifying, and do not become fully ohmic even after annealing at 780 °C.

The current-voltage trace of the preliminary p - n device is shown in Fig. 8. It shows roughly diodelike behavior, although the nonideal effects of high series resistance and rectifying contacts are evident. Turn on voltage is seen to be ~ 3 V, which is expected, as it corresponds to the band gap of the materials. The necessity of lateral contacts and the low conductivity of GaN are largely responsible for the large series resistive component.

At low applied voltages (3–10 V), luminescence is observed by eye to be a pale blue similar to the hue observed in ZnS/ n -GaN PL [Fig. 6(B)]. The devices are bright enough during room temperature operation to be seen with the room lights on with 1 mA current applied. The expected nonuniformity of luminescence due to current spreading is observed. Additionally, "bright spots" are seen which may be caused by defects or thickness variations in the ZnS layer. As forward bias is increased, light appears more violet colored, similar to low temperature ZnS/ p -GaN PL [Fig. 6(A)]. The

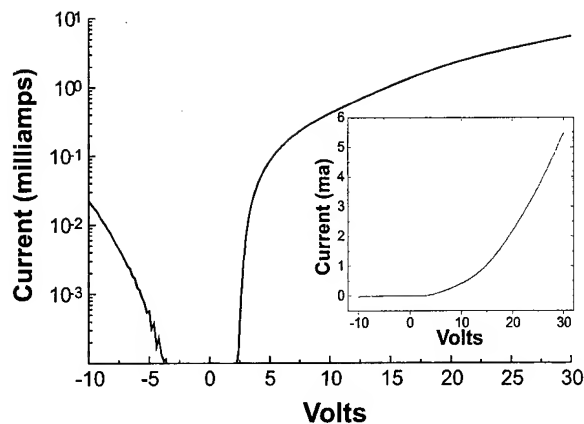


FIG. 8. Current-voltage trace of the prototype p -GaN/ n -ZnS light-emitting device. Diode-like behavior is observed with large series resistance effects.

exact mechanism of luminescence is unclear at this time, as there are several possibilities. The rather broad photoluminescence spectra of the GaN:Mg and ZnS:Al,Ag samples were found to be quite similar at room temperature, thus making it difficult to deduce if recombination is taking place in the GaN layer or the ZnS layer. Other mechanisms, such as recombination in the nitride layer followed by absorption and fluorescence in the sulfide layer, are possible. Also, luminescence at high biases may be linked to hot carrier excitation processes. Due to the low temperatures of growth and device processing, we do not expect that a significant amount of interdiffusion has occurred, such as Zn diffusing into the nitride layer or Ga diffusing into the sulfide layer. Ga is known to act as a donor in ZnS, and Zn a deep acceptor in GaN. Spectroscopy and other work is in progress in the study of this structure.

V. SUMMARY

Zinc sulfide thin films were grown on GaN and sapphire substrates by MBE, with the GaN surface being effectively cleaned prior to growth by exposure to a rf nitrogen plasma. RHEED patterns indicated that monocrystalline films could be grown on both substrates, but most contained microtwins in the (111) or ($\bar{1}\bar{1}\bar{1}$) planes. No twins were visible in the films grown at 400 °C. Typical growth rates were measured at 0.1–0.2 $\mu\text{m/h}$ for temperatures below 350 °C and were much lower for higher temperatures. The samples were characterized using photoluminescence and x-ray diffraction. X-ray peaks typically had FWHM of 400 arcsec for $\omega/2\theta$ scans, and larger widths for ω scans, indicating a degree of mosaic tilts. The PL from the ZnS films doped with Ag and Al displayed bright blue luminescence. Heterostructure devices were fabricated and their current-voltage curves were measured. Electroluminescence from these devices was pale blue at low forward bias and shifted to violet at higher biases.

ACKNOWLEDGMENTS

This work was supported by Advanced Research Project Agency, and monitored by Office of Naval Research under Grant No. N00014-92-J-1845.

- ¹C. N. King, J. Vac. Sci. Technol. A **14**, 1729 (1996).
- ²J. L. Castellano, *Handbook of Display Technology* (Academic, San Diego, CA, 1992), Chap. 6.
- ³Y. Sato, N. Takahashi, and S. Sato, Jpn. J. Appl. Phys. **1** **351**, L838 (1996).
- ⁴S. Yamaga, Physica B **185**, 500 (1993).
- ⁵B. J. Wu, L. H. Kuo, J. M. Depuydt, G. M. Haugen, M. A. Haase, and L. Salamancariba, Appl. Phys. Lett. **68**, 379 (1996).
- ⁶S. Nakamura, N. Iwasa, M. Senoh, and T. Mukai, Jpn. J. Appl. Phys. **1** **131**, 1258 (1992).
- ⁷H. Amano, M. Kito, K. Hiramatsu, and I. Akasaki, Jpn. J. Appl. Phys. **1** **128**, L2112 (1989).
- ⁸C. Yuan, T. Salagaj, A. Gurary, P. Zawadzki, C. S. Chern, W. Kroll, and R. A. Stall, J. Electrochem. Soc. **142**, L163 (1995).
- ⁹M. W. Wang, J. O. McCaldin, J. F. Swenberg, T. C. McGill, and R. J. Hauenstein, Appl. Phys. Lett. **66**, 1974 (1995).
- ¹⁰J. O. McCaldin, M. W. Wang, and T. C. McGill, J. Cryst. Growth **159**, 502 (1996).
- ¹¹M. Yokoyama, K. Kashiro, and S. Ohta, J. Cryst. Growth **81**, 73 (1987).
- ¹²I. P. McClean and C. B. Thomas, Semicond. Sci. Technol. **7**, 1394 (1992).
- ¹³T. Yao and S. Maekawa, J. Cryst. Growth **53**, 423 (1981).
- ¹⁴J. W. Cook, Jr., D. B. Eason, R. P. Vaudo, and J. F. Schetzina, J. Vac. Sci. Technol. B **10**, 901 (1992).
- ¹⁵K. B. Ozanyan, L. May, J. E. Nicholls, J. H. C. Hogg, W. E. Hagston, B. Lunn, and D. E. Ashenford, Solid State Commun. **97**, 345 (1996).
- ¹⁶O. Kanehisa, M. Shiiki, M. Migita, and H. Yamamoto, J. Cryst. Growth **86**, 367 (1988).
- ¹⁷S. Ohta, K. Kashiro, and M. Yokoyama, J. Cryst. Growth **87**, 217 (1988).
- ¹⁸K. Yoneda, T. Toda, Y. Hishida, and T. Niina, J. Cryst. Growth **67**, 125 (1984).
- ¹⁹J. E. Ayers, S. K. Ghandhi, and L. J. Schowalter, J. Cryst. Growth **113**, 2156 (1991).
- ²⁰K. Ichino, T. Onishi, Y. Kawakami, S. Fujita, and S. Fujita, J. Cryst. Growth **138**, 28 (1994) (see discussion section).
- ²¹Z. Z. Bandić, E. C. Piquette, J. O. McCaldin, and T. C. McGill (unpublished).
- ²²*Physics and Chemistry of II-VI Compounds*, edited by M. Aven and J. S. Prener (North-Holland, Amsterdam, 1967), Chap. 9.

Controlled formation of organic layers on semiconductor surfaces

Jennifer S. Hovis, Seung Lee, Hongbing Liu, and Robert J. Hamers^{a)}
Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706

(Received 12 January 1997; accepted 28 April 1997)

It has been shown that well-defined, ordered organic layers can be formed on the silicon (100) surface. This is achieved through the interaction of unsaturated C=C bonds with the oriented dimers of the reconstructed Si(100)-(2×1) surface. In this article, we present an investigation of the structure and chemical bonding of organic films prepared using different organic precursors. Data were obtained using scanning tunneling microscopy, Fourier-transform infrared spectroscopy, and x-ray photoelectron spectroscopy. The molecules investigated are cyclopentene, 3-pyrroline, and norbornadiene, representing prototypical cyclic, heterocyclic, and bicyclic unsaturated organic molecules, respectively. Each molecule has at least one unsaturated C=C bond. © 1997 American Vacuum Society. [S0734-211X(97)11804-2]

I. INTRODUCTION

In recent years, there has been a great deal of interest in organic-based materials for applications such as nonlinear optics, thin-film displays, lithography, and molecular electronics. The controlled integration of organic materials with the existing infrastructure in silicon-based microelectronics technology would present many new opportunities in these areas.

Chidsey and co-workers¹ successfully formed layers of densely packed alkyl monolayers on hydrogen-terminated Si(111) and Si(100) surfaces through a free-radical mechanism. Also, ethylene and acetylene have both been shown to be capable of bonding to the Si(100) surface through a so-called "di- σ " configuration. In this configuration, the double bonds of the Si dimers and the double bonds of the adsorbed species are broken, forming two new Si-C σ bonds.²⁻⁹

In an earlier report,¹⁰ we showed that cyclopentene interacted with the dimers of Si(100) leading to well-defined monolayers with the molecules in the organic layer exhibiting both translational and rotational order over macroscopic dimensions. The basis of our method is that the dimers (formally two Si atoms sharing a σ and a π bond), which comprise the clean Si(100)-(2×1) surface can interact with double bonds in unsaturated organic molecules, thus, attaching the molecule to the surface through the formation of two new Si-C bonds. Because the interacting bonds are both strongly oriented, the interaction of the Si=Si dimers with the unsaturated C=C bonds effectively makes the surface Si=Si dimers act as a template for extending rotational order into the organic film.

It should be possible to generalize this methodology to other unsaturated organic molecules so as to prepare a variety of organic layers having novel physical and/or chemical properties. In this article, we compare the interaction of three different unsaturated organics with Si(100). Cyclopentene, which is shown interacting with the silicon surface in Fig. 1(a) is a cyclic molecule with one unsaturated C=C unit.

3-pyrroline, Fig. 1(b), is a heterocyclic molecule having an amine group as part of the ring structure. Because there are a number of well-defined organic chemistry reactions for linking secondary amines to other chemical functionalities, such a molecule can serve as a model for exploring the possibility of generalized attachment chemistry on the Si(100) surface. Finally, norbornadiene, Fig. 1(c) is a bicyclic molecule having two unsaturated bonds.

II. EXPERIMENT

All experiments reported here were performed in ultra-high vacuum (UHV) systems achieving base pressures of $< 1 \times 10^{-10}$ Torr. The surfaces were characterized by several techniques. Direct observation of adsorbed molecules and molecular layers was achieved using a home-built UHV scanning tunneling microscope (STM). All STM images shown here were obtained with a sample bias of -3 V and a tunneling current of 0.2-0.3 nA. Fourier-transform infrared (FTIR) adsorption spectra were obtained using a multiple internal-reflection geometry and a Mattson RS-1 FTIR spectrometer coupled to a UHV system through BaF₂ windows. X-ray photoelectron spectroscopy (XPS) experiments were performed using a Physical Electronics system with a monochromatized Al K α source. For each type of experiment, samples were prepared and characterized completely *in situ*.

The molecules used for growing the films reported on were bought from Aldrich and each had a purity of at least 97%. They were introduced into the chamber through a variable leak valve. Clean Si(100) samples were prepared by annealing pieces of Si(100) wafers to 1475 K while maintaining the chamber pressure at $< 3 \times 10^{-10}$ Torr. This procedure produces a clean well-ordered surface exhibiting a (2×1) reconstruction.¹¹ Samples having two different levels of doping and two slightly different crystallographic orientations were used. Highly doped ($< 0.15 \Omega$ cm resistivity, Sb- and P-doped) samples were used for STM and XPS experiments, while lightly doped ($> 5 \Omega$ cm resistivity, P-doped) samples were used for infrared spectroscopy experiments to reduce free-carrier adsorption in the bulk.

^{a)} Author to whom correspondence should be addressed; Electronic mail: rjhamers@facstaff.wisc.edu

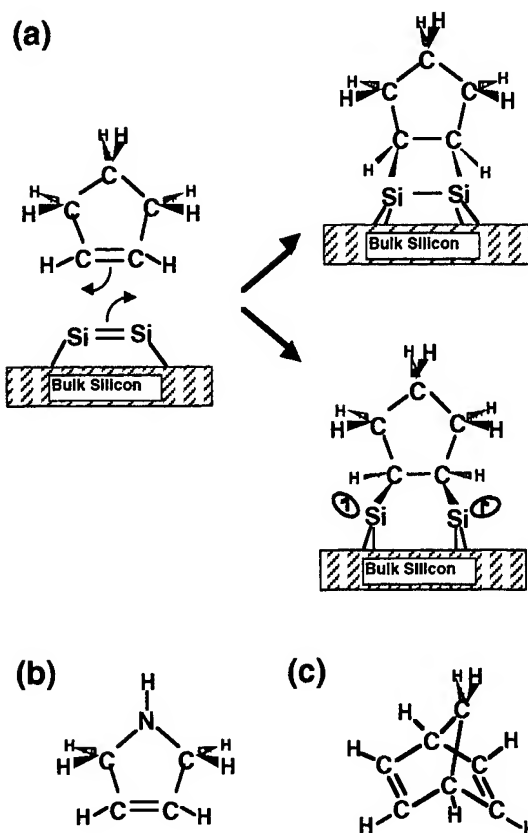


FIG. 1. Schematic illustration of molecules discussed in the text. (a) Interaction of cyclopentene with the Si(100) surface. (b) 3-pyrroline. (c) Norbornadiene.

Two slightly different crystallographic orientations of Si(100) were used. On-axis samples were oriented to expose the (100) surface with an accuracy $\pm 0.5^\circ$. The distance between surface steps is large for on-axis wafers. As a result, the steps are only one atom high, and the direction of the Si=Si dimer bond rotates by 90° across each step; this leads to a surface containing equal amounts of the two different rotational domains of the dimers, with each domain separated by a single-height step. We also performed experiments using samples purposely miscut by $4.0^\circ \pm 0.5^\circ$ toward the $\langle 110 \rangle$ direction. For such 4° -miscut off-axis ("vicinal") samples, the interactions between adjacent steps are sufficiently strong that almost all steps on the surface are two atoms high.^{12,13} As a result, the rotational orientation of the Si=Si dimers is retained across the step, leading to samples in which there is only one rotational domain of the Si=Si dimers. In the FTIR experiments, we utilized several different polarizations of the incident electric field. In particular, we note that for *s*-polarized light (in which the electric field is parallel to the surface plane) the interaction of the electric field with the Si=Si dimers on a vicinal surface is different, depending on whether the electric field is parallel to or perpendicular to the Si=Si dimer axis. Figure 2 schematically depicts the direction of the electric field relative to the silicon dimers for two different polarizations. Polarization directions

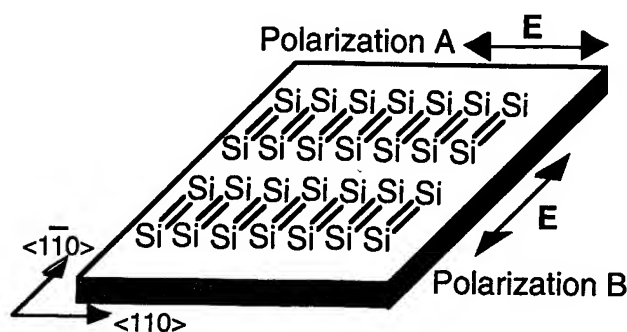


FIG. 2. Schematic illustration of the two polarization directions used in FTIR experiments.

"A" and "B" both correspond to *s*-polarized light, but direction "A" has the electric field vector E oriented along the $\langle 110 \rangle$ direction (perpendicular to the Si=Si dimer bonds) and polarization "B" has E oriented along the $\langle 1\bar{1}0 \rangle$ direction (parallel to the Si=Si dimer bonds).

III. RESULTS AND DISCUSSION

A. Films formed from exposure of Si(100) to cyclopentene

In Fig. 3, we show scanning tunneling microscopy images after Si(100) was exposed to 30 L (1×10^{-7} Torr for 300 s) of cyclopentene at room temperature. In Figs. 3(a) and 3(b), the surface is oriented to within 0.5° of the (100) plane, in Fig. 3(c), the surface is a 4° -miscut vicinal surface, with arrows pointing to the locations of the double-height steps. In all three images, each oval shaped object protruding from the surface is an individual cyclopentene molecule, with the direction of elongation corresponding to the wide direction of the ring. In Fig. 3(c), it can be seen that the orientation of the molecules is retained across the double-height steps; additionally, the translational ordering in Fig. 3(c) is somewhat better than that observed for the on-axis sample, Figs. 3(a) and 3(b). We believe the improvement in ordering is due to the fact that the steps, which are separated by only ~ 20 Å, provide a means for relieving strain, which builds up in the organic layer due to steric repulsion between adjacent molecules. Regardless of any imperfections in the translational ordering of the films, there are two very important features that can be discerned in all three images. First, the molecules are clearly aligned into rows; this demonstrates that the cyclopentene molecules spontaneously order with very specific bonding locations. Second, the individual molecules appear elongated, and this direction of elongation is the same for all molecules. These images, therefore, prove that the molecules are ordered rotationally as well as translationally. The use of 4° -miscut Si(100) surfaces, Fig. 3(c), allows the rotational orientation to be retained across the entire sample (length scale of centimeters).^{12,13}

The rotational anisotropy in the molecular orientation should yield rotational anisotropy in optical properties. To demonstrate that this is indeed the case, we present infrared

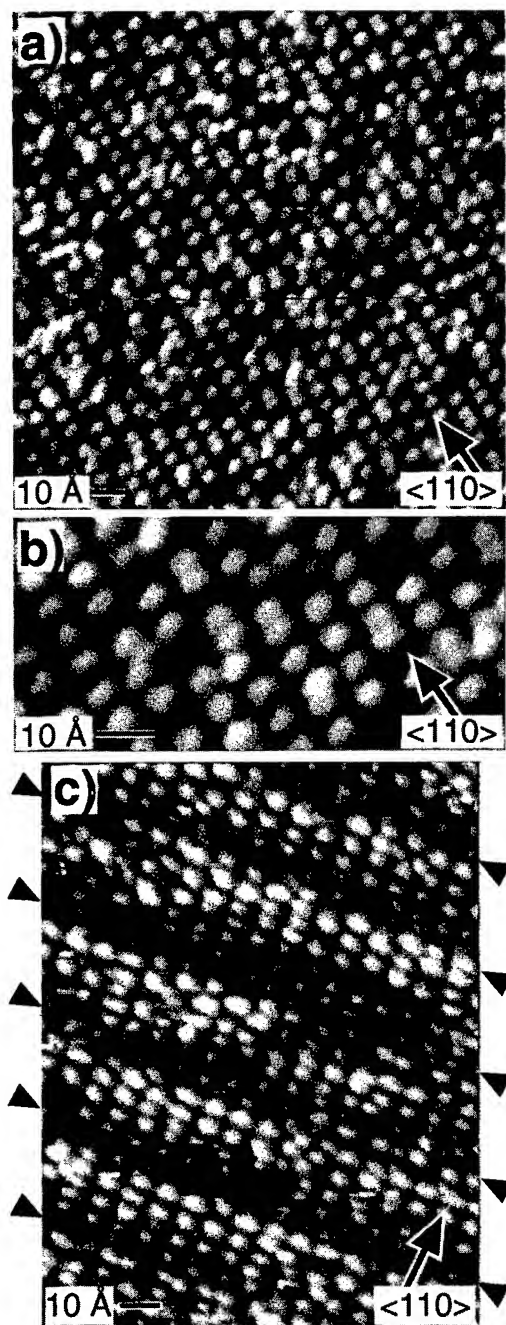


FIG. 3. STM image of Si(100) surface after exposure to 30 L of cyclopentene at room temperature, bias -3 V, tunneling current 0.2 – 0.3 nA. (a) Image showing ordered molecular adsorption on a single terrace. (b) High-resolution image showing the translational ordering and consistent elongated appearance of individual molecules. (c) Image obtained on a 4° -miscut vicinal Si(100) substrate showing that molecular orientation of cyclopentene is preserved across double-height steps. Arrows indicate the positions of the double-height steps.

absorption spectra for vicinal Si(100) samples exposed to cyclopentene. Figure 4 shows infrared absorption spectra for s -polarized light parallel to and perpendicular to the direction of the Si=Si dimer bonds. As shown in Fig. 4, the spectra show significant differences between the two polarization directions in the C–H stretching region

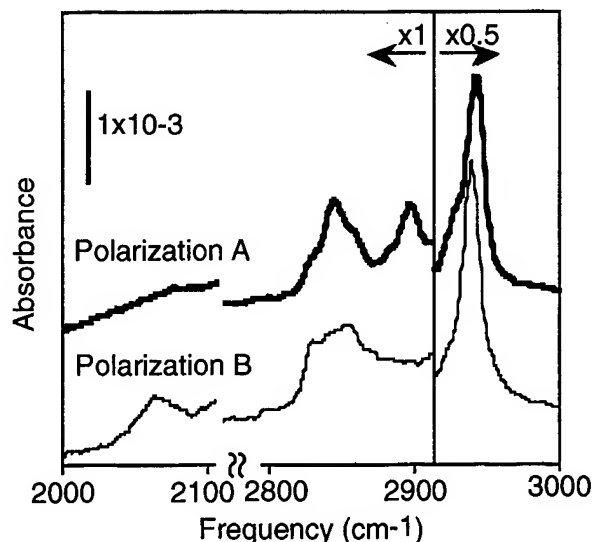


FIG. 4. Si–H and C–H spectral regions showing the anisotropy of infrared absorption of Si(100) after exposure to 30 L at room temperature of cyclopentene for both polarization directions of s -polarized light.

(2800 – 3000 cm^{-1}). In particular, the absorption peak at 2895 cm^{-1} is several times larger when measured using polarization direction “A” than with direction “B.” Differences are also observed in the relative intensities of the overlapping peaks at 2850 – 2875 cm^{-1} . Although a full spectral analysis has not yet been performed, we attribute the peak at 2895 cm^{-1} to a mode involving motion of the two H atoms at the apex of the cyclopentene molecule (attached to the C atom farthest from the Si surface), these C–H bonds are expected to lie in a plane perpendicular to polarization direction “B” but parallel to polarization direction “A.” All other C–H bonds in this molecule are expected to have components of their dynamic dipole moment along both $\langle 110 \rangle$ and $\langle \bar{1}\bar{1}0 \rangle$ directions, meaning they should exhibit some absorption strength for both polarizations. The fact that differences in absorption for s -polarized light along these two directions is seen confirms that the orientation observed in the STM image, Fig. 3(c), is maintained over centimeter distances.

The Si–H stretching region in Fig. 4 shows a small peak near 2070 cm^{-1} for polarization “B.” This peak corresponds to the Si–H stretching mode for Si dimers with one attached hydrogen atom,^{14,15} and suggests that there is some dissociation on the surface. It is not yet clear whether the dissociation arises from impurities in the cyclopentene or if it is intrinsic to the preparation procedure. Nevertheless, it is obvious from the small size of this peak and the ordering observed in Fig. 3 that the extent of dissociation is small (we estimate $<5\%$), and does not significantly affect the ordering of the molecular film.

To confirm that the cyclopentene molecules are bonded as depicted in Fig. 1, we obtained x-ray photoelectron spectroscopy data for these films, as shown in Fig. 5. Figure 5 shows two C $1s$ peaks having binding energies of 284.9 and 284.2 eV with an intensity ratio $3:2$. Since C is more electronega-

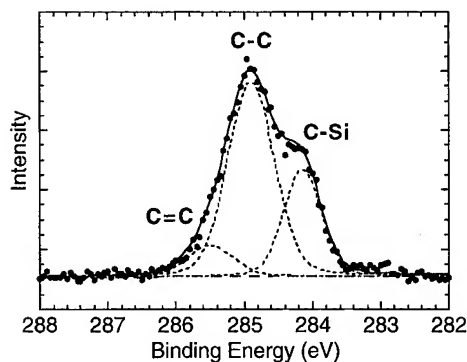


FIG. 5. C 1s XPS spectra from a Si(100) surface, which was exposed to 30 L of cyclopentene at room temperature.

tive than Si, the C atoms bonded directly to the Si surface are expected to be more negatively charged and to, thus, have smaller core-level binding energies than the C atoms of the cyclopentene molecule, which are not bonded to the surface. We, therefore, attribute the peak at 284.2 eV to the two C atoms bonded directly to the Si surface, and the more intense peak at 284.9 eV to the other three C atoms having no direct C–Si bond. There is also a small component of the C 1s peak at a binding energy of 285.5 eV. This is due to C=C and arises from excess cyclopentene physisorbed to the surface.

One remaining question is whether the Si–Si bond of the silicon dimers remains intact. Two possibilities exist, as depicted in Fig. 1(a); if the bond remains intact, a highly strained four-member ring involving two Si and two C atoms is formed. While the thermal 2+2 cycloaddition reaction is formally disallowed by symmetry arguments, dimer tilting might destroy the symmetry and allow the thermal reaction to take place. However, we note that previous studies of ethylene on Si(100) have suggested that the Si–Si dimer bond might be broken, leading to a larger ring structure and Si atoms having formal “dangling bonds.”^{8,9} None of the experiments we have done, to date, have allowed us to conclusively resolve this question.

B. Films formed from the exposure of Si(100) to 3-pyrroline

While cyclopentene shows a high degree of translational and rotational order, further chemical functionalization is difficult. However, the presence of another functional group within the ring structure, such as an amine group, should provide a means for further chemical functionalization of the organic film and/or a means for continuing the controlled growth into subsequent layers to produce multilayer films.

The molecule 3-pyrroline, as shown in Fig. 1(b), is perhaps the simplest modification of the cyclopentene molecule. In Fig. 6, we show scanning tunneling microscopy images after Si(100) was exposed to 30 L of 3-pyrroline at room temperature. Each protrusion **P** in the images corresponds to an individual 3-pyrroline molecule bonding to the surface. Once again, high-resolution STM images show that these **P**



FIG. 6. STM image of the Si(100) surface after exposure to 30 L of 3-pyrroline at room temperature, bias -3 V, tunneling current 0.2 nA.

protrusions are slightly elongated parallel to the dimer direction, suggesting that the molecules bond in a manner similar to cyclopentene. While the ordering for 3-pyrroline is poorer than for cyclopentene, it is important to note that there exist areas showing both translational and rotational order, demonstrating that there is a preferred bonding of 3-pyrroline at specific surface sites. However, the relative disorder and analysis of the positions of the features observed in STM suggest that there is more than one way for 3-pyrroline to bond to the Si(100) surface.

The presence of more than one bonding configuration is confirmed through x-ray photoelectron spectra. Figure 7(a) shows N 1s XPS spectra taken of a Si(100) sample that was exposed to 1 L of 3-pyrroline at room temperature and then, subsequently, annealed to successively higher temperatures for 1 min. Immediately after exposure of Si(100) to 3-pyrroline at room temperature, two separate N 1s peaks of roughly equal intensity are observed having binding energies of 398.4 and 399.4 eV. As the sample is subsequently annealed to higher temperatures, the peak at low binding energy decreases in intensity, while the peak at higher binding energy remains intact with constant intensity. After annealing to 500 °C, only a single N 1s peak is observed at approximately 399.4–399.6 eV. The shift in this peak as the temperature increases is due to band bending, the silicon 2p peak shows the same shifting as the temperature increases. At higher temperatures, this peak disappears and is replaced by a new peak at 397.5 eV; this peak persists to very high temperatures and is attributed to nitrogen incorporation into the silicon giving rise to silicon nitride. Based on this information, a logical assignment of the 3-pyrroline bonding configurations giving rise to the two N 1s peaks is as follows: the 398.4 eV peak is due to the N atom bonding to the Si surface directly, through the formation of a N–Si bond, while the 399.4 eV peak is attributed to bonding through an

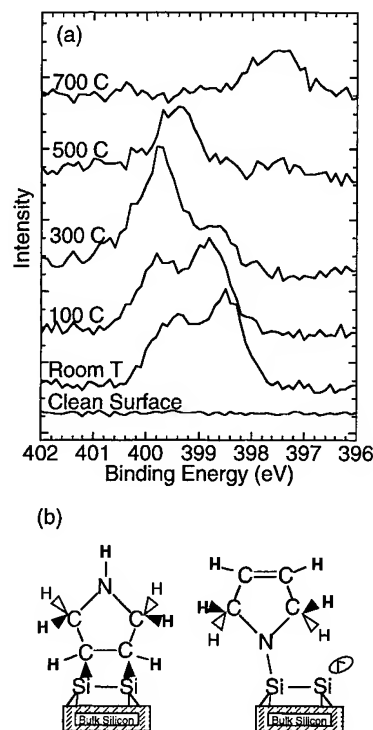


FIG. 7. (a) N 1s XPS spectra taken of a Si(100) sample that was exposed to 1 L of 3-pyrroline at room temperature and then subsequently annealed to successively higher temperatures for 1 min. (b) Schematic illustration of the two different bonding possibilities for 3-pyrroline.

addition/insertion reaction similar to that observed for cyclopentene. These bonding configurations are shown in Fig. 7(b). This assignment is consistent with the fact that Si is less electronegative than H, thus, a N atom bonded to Si would be expected to have a smaller core-level binding energy than a N atom bonded to H. The disappearance of the 399.4 eV peak occurs at a relatively high temperature that is similar to the temperature at which ethylene desorbs from Si(100).^{6,8} Ethylene and 3-pyrroline (in the bonding configuration assigned to this peak) are both bonded to the surface through two Si-C bonds.

In Fig. 8, we show FTIR spectra, using polarization direction "A," of Si(100) after exposure to 8 L of 3-pyrroline at room temperature. While a full spectral analysis is impossible at this point, there are still two useful observations we can make. First, there is a peak near 2075 cm^{-1} , which is in the Si-H stretching region. This peak is due to a hydrogen atom adsorbed to a silicon dimer,^{14,15} and its presence indicates that there is some dissociation upon adsorption, leading to surface hydrogen. Also, there is a small peak near 3073 cm^{-1} , which is the spectral range attributed to the C-H stretching mode of an alkene. This peak is attributed to C-H stretches of 3-pyrroline molecules that are bonded to the Si surface through the N atom; presumably, this bonding occurs with concurrent loss of the amine H atom, which then becomes bonded to the underlying Si surface. Thus, the 3073 cm^{-1} peak and the 2075 cm^{-1} peak result from bonding of the 3-pyrroline molecule to the Si surface directly through the N atom. Analysis of the C-H region has not

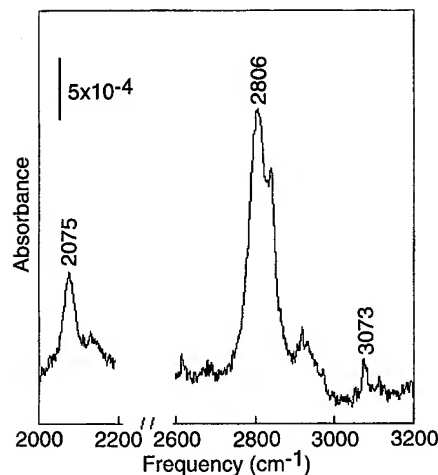


FIG. 8. FTIR spectra of polarization direction "A" of Si(100) after exposure to 8 L of 3-pyrroline at room temperature.

been successful due to the presence of a very broad peak over 100 cm^{-1} wide.

C. Films formed from the exposure of Si(100) to norbornadiene

Norbornadiene, Fig. 1(c), was investigated to see if the methodology described above could be extended to molecules having more than one unsaturated double bond. Figure 9 shows FTIR data of both polarization directions "A" and "B" after vicinal Si(100) was exposed to 10 L of norbornadiene at room temperature. The first thing to notice about these spectra is that there are no visible peaks in the Si-H stretching region, but there are peaks in the C-H stretching region. This immediately demonstrates that norbornadiene adsorbs on the surface with little or no dissociation. In order for norbornadiene to interact with the surface without a loss of hydrogen, it must do it through the unsaturated C=C bond. Second, there is a relatively strong peak at 3057 cm^{-1} , which is characteristic of the C-H stretch of

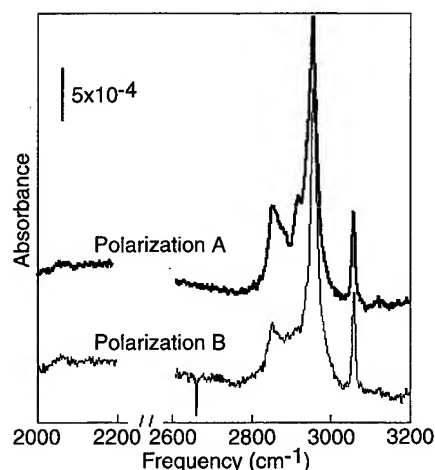


FIG. 9. FTIR spectra using both polarization directions of s light of (100) after exposure to 10 L of norbornadiene at room temperature.

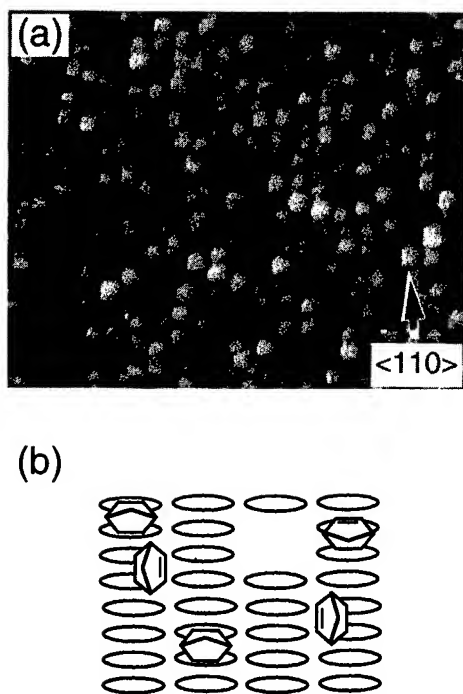


FIG. 10. (a) STM image of Si(100) surface after exposure to 30 L of norbornadiene at room temperature, bias -3.6 V, tunneling current 0.3 nA. (b) Schematic illustration of several bonding sites for norbornadiene on Si(100).

an unsaturated molecule. This, in turn, suggests that in at least some of the molecules, only one (or none) of the two unsaturated $C=C$ bonds have been eliminated by bonding to the Si(100) surface. Finally, a closer comparison of the two different polarization directions shows that there is little difference between them. This indicates that while we know that the molecule is interacting with the surface through the $C=C$ bond, there must be more than one type of surface site at which the interaction can occur.

Scanning tunneling microscopy images, as shown in Fig. 10(a), give further insight into how norbornadiene bonds to the Si(100) surface. The STM data were taken after vicinal Si(100) was exposed to 30 L of norbornadiene at room temperature. There is no translational or rotational ordering apparent in this image, and it is difficult to see where the double-height steps are. The absence of Si-H vibrations in the FTIR data strongly suggests that cleavage of C-H bonds does not occur, leading us to conclude that each protrusion observed in Fig. 10(a) is a single norbornadiene molecule. In studies performed at lower coverage (not shown), we find that some of the features attributed to norbornadiene appear to be on top of the Si=Si dimers and some appear to be bonded midway between the rows of Si=Si dimers. Therefore, it is reasonable to assume that the lack of order in the film and the absence of anisotropy in the FTIR spectra is due to the molecule bonding to the surface in more than one configuration. Figure 10(b) shows four bonding possibilities: both $C=C$ bonding within a dimer row, one $C=C$ bonding within a dimer row, both $C=C$ bonding across a dimer row, and one $C=C$ bonding across a dimer row. Analysis of the

resulting FTIR spectra is, therefore, complicated due to the 90° rotation of some molecules with respect to each other.

While the interaction of norbornadiene with Si(100) results in a mixture of several different bonding geometries, it should be noted that the organic film shows virtually no dissociation. Thus, this molecule, and/or derivatives of it, may still serve as useful interfacial precursors for the creation of organic thin films.

IV. SUMMARY AND CONCLUSIONS

We have identified a general class of reactions involving the interaction of unsaturated organic molecules with the Si(100) surface that enable organic monolayers to be prepared on Si(100) in a controlled manner. The three molecules presented in this article grow films through the same process but with different net results. Cyclopentene can be used to grow monolayer films with a high degree of translation and rotational ordering. The films grown from the heterocyclic molecule 3-pyrroline have areas showing both translational and rotational ordering. While, for this molecule, there are two possible bonding configurations to the Si surface at room temperature, control of temperature and/or other experimental variables may permit optimization of one particular structure. Finally, the bicyclic molecule norbornadiene leads to layers that appear disordered in STM due to more than one bonding configuration; however, these monolayer films show virtually no cleavage of C-H bonds and apparently bond to the Si surface through the $C=C$ π system in a manner similar to the other molecules.

ACKNOWLEDGMENTS

This work was supported by the U.S. Office of Naval Research and the National Science Foundation Presidential Faculty Fellowship Program.

- ¹M. R. Linford, P. Fenter, P. M. Eisenberger, and C. E. D. Chidsey, *J. Am. Chem. Soc.* **117**, 3145 (1995).
- ²C. C. Cheng, R. M. Wallace, P. A. Taylor, W. J. Choyke, and J. T. Yates, Jr., *J. Appl. Phys.* **67**, 3693 (1990).
- ³P. A. Taylor, R. M. Wallace, C. C. Cheng, W. H. Weinberg, M. J. Dresser, W. J. Choyke, and J. T. Yates, Jr., *J. Am. Chem. Soc.* **114**, 6754 (1992).
- ⁴A. J. Mayne, T. R. I. Cataldi, J. Knall, A. R. Avery, T. S. Jones, L. Pinheiro, H. A. Hill, G. A. D. Briggs, J. B. Pethica, and W. H. Weinberg, *Faraday Discuss. R. Soc.* **94**, 199 (1992).
- ⁵A. J. Mayne, A. R. Avery, J. Knall, T. S. Jones, G. A. D. Briggs, and W. H. Weinberg, *Surf. Sci.* **284**, 247 (1993).
- ⁶L. Clemen, R. M. Wallace, P. A. Taylor, M. J. Dresser, W. J. Choyke, W. H. Weinberg, and J. J. T. Yates, *Surf. Sci.* **92**, 205 (1992).
- ⁷C. Huang, W. Widdra, and W. H. Weinberg, *Surf. Sci. Lett.* **315**, 953 (1994).
- ⁸W. Widdra, C. Huang, and W. H. Weinberg, *Surf. Sci.* **329**, 295 (1995).
- ⁹W. Widdra, C. Huang, S. I. Yi, and W. H. Weinberg, *J. Chem. Phys.* **105**, 5605 (1996).
- ¹⁰R. J. Hamers, J. Hovis, S. Lee, H. Liu, and J. Shan, *J. Phys. Chem.* **101**, 1489 (1997).
- ¹¹R. J. Hamers, R. M. Tromp, and J. E. Demuth, *Phys. Rev. B* **34**, 5343 (1986).
- ¹²J. D. Chadi, *Phys. Rev. Lett.* **59**, 1691 (1987).
- ¹³M. Henzler and J. Clabes, *Jpn. J. Appl. Phys. Suppl. Pt. 2* **2**, 389 (1974).
- ¹⁴J. Shan, Y. Wang, and R. J. Hamers, *J. Phys. Chem.* **100**, 4961 (1995).
- ¹⁵Y. J. Chabal and K. Raghavachari, *Phys. Rev. Lett.* **53**, 282 (1984).

Gallium desorption behavior at AlGaAs/GaAs heterointerfaces during high-temperature molecular beam epitaxy

K. Mahalingam and D. L. Dorsey

Materials Directorate, Wright Laboratory (WL/MLPO), Wright-Patterson Air Force Base, Ohio 45433-7707

K. R. Evans^{a)}

Avionics Directorate, Wright Laboratory (WL/AADP), Wright-Patterson Air Force Base, Ohio 45433-7707

Rama Venkat

Department of Electrical Engineering, University of Las Vegas, Las Vegas, Nevada 89154

(Received 12 January 1997; accepted 22 April 1997)

A Monte Carlo simulation study is performed to investigate the Ga desorption behavior during AlGaAs-on-GaAs heterointerface formation by molecular beam epitaxy. The transients in the Ga desorption rate upon opening the Al shutter are shown to be associated with the concurrent reduction in the V/III flux ratio. Monte Carlo simulations employing a constant V/III flux ratio yield a "steplike" variation in the Ga desorption rate with the resulting interfaces closer in abruptness to the ideal AlGaAs-on-GaAs interface. Further details on the stoichiometry of the interface and its relationship with predicted Ga desorption profiles is presented. © 1997 American Vacuum Society. [S0734-211X(97)09704-7]

I. INTRODUCTION

AlGaAs/GaAs heterostructures grown by molecular beam epitaxy (MBE) are important materials for applications in optical and electronic devices. It has been established for this system that growth at a high substrate temperature has important advantages. In addition to producing interfaces of high structural quality, employing high substrate temperatures also improves the optoelectronic properties of AlGaAs.^{1,2} An important problem that arises, however, is that these substrate temperatures are high enough to cause significant Ga desorption.³⁻⁷ Furthermore, *in situ* desorption mass spectrometry studies^{8,9} reveal that during AlGaAs-on-GaAs heterointerface formation the Ga desorption behavior is more complex than a simple constant or "steplike" behavior. Upon opening the Al shutter, there is a sharp rise in the Ga desorption rate, followed by a gradual decrease to a steady state value less than that corresponding to GaAs growth. A potential consequence of this transient is a compositional grading that extends 2-4 monolayers (ML) into the AlGaAs layer.

In a recent study¹⁰ we developed a Monte Carlo (MC) model to investigate the possible mechanisms that could explain the observed Ga desorption behavior at the AlGaAs-on-GaAs interface as seen during desorption mass spectrometry (DMS) experiments. The various factors investigated include the Al-Ga surface exchange mechanism, the effects of change in V/III flux ratio accompanying the opening of the Al shutter, and the influence of Al-Ga interaction energy. It was observed that the model including the Al-Ga surface exchange mechanism agreed well with the experimental results. Interesting was the fact that it was also observed that the experimental results could be reproduced

even without the Al-Ga surface exchange mechanism when the Al-Ga interaction energy was set to be equal to the Ga-Ga interaction energy. An examination of the predicted time evolution of the surface stoichiometry revealed that the transients in the Ga desorption rate were associated with a sharp decrease in the As coverage due to a reduction in the V/III ratio accompanying the opening of the Al shutter (to be discussed further later). The sharp rise in the Ga desorption rate was explained in terms of an instantaneous increase in the surface population of weakly bound Ga species due to the reduction in As coverage. The subsequent gradual decrease in the Ga desorption ratio was explained by the reduction in the number of surface cation sites from which Ga could desorb due to the presence of Al.

The above results indicate that the key to eliminating the Ga desorption transients at the AlGaAs-on-GaAs heterointerface is to maintain a constant V/III flux ratio. In this study we extend previously developed MC models¹⁰ to investigate different strategies that employ a constant V/III flux ratio. The effectiveness of the different strategies is examined by comparing the predicted Al/Ga concentration profiles at the AlGaAs-on-GaAs interface.

II. MODEL

The MC simulation procedure employed in the present study is similar to that adopted in previous models,^{11,12} wherein MBE growth is simulated by including deposition, surface migration, and desorption as allowed events. The model simulates growth on an unreconstructed (100) zinc-blende surface, explicitly taking into account the tetrahedral nature of the lattice with its individual sublattices for the cations and anions. An important difference between the present and previous models is that in the present model a temperature-dependent physisorption mechanism for cations

^{a)}Present address: Quantum Epitaxial Designs, Inc., 119, Technology Drive, Bethlehem, PA 18015.

is included to accurately model their desorption kinetics at high growth temperatures. In this mechanism a cation arriving at a cation terminated site is allowed a fixed number (temperature dependent) of jumps, determined by fit to experimental data, to find an anion terminated site; the cation desorbs if it is unsuccessful. The significance of this mechanism is that it accounts for the finite surface residence time for cations at high temperatures. By including this mechanism it was possible to obtain quantitative agreement with experimental Ga desorption rates for a range of growth temperatures,⁸ and also to accurately predict the dependence of Ga desorption energy on V/III flux ratio (for GaAs homoepitaxy) observed in experiments.⁵

Growth simulations were performed on a substrate 40×40 in size (site ²) assuming a solid-on-solid (SOS) model (vacancies and overhangs not permitted) with periodic boundary conditions. The kinetic rates for deposition of Al, Ga, and As were assumed to be equal to their respective fluxes. The rates for surface migration and desorption events were assumed to be of the Arrhenius form with configuration-dependent activation energies calculated in terms of the first- and second-nearest-neighbor interaction energies, as in earlier MC models.^{11,12} The frequency factors and interaction energies for GaAs were taken from a previous study,¹⁰ and those for AlAs were determined from values for the cohesive energy.¹³ The model parameters used were the frequency factors for diffusion and desorption, $R_{\text{diff}} = 10^{10}/\text{s}$, $R_{\text{des}} = 7.5 \times 10^{12}/\text{s}$, respectively; nearest-neighbor interaction energies, $E_{\text{Ga-As}} = 0.85$ eV, $E_{\text{Al-As}} = 0.97$ eV; and second-nearest-neighbor interaction energies $E_{\text{Ga-Ga}} = 0.17$ eV, $E_{\text{As-As}} = 0.12$ eV, and $E_{\text{Al-Al}} = 0.25$ eV.

Our previous MC simulation study¹⁰ has shown that the Ga desorption behavior at the AlGaAs-on-GaAs interface is best described by two models (hereafter referred to as models I and II), where model I includes the Al-Ga surface exchange mechanism with the Al-Ga interaction energy ($E_{\text{Al-Ga}} = 0.21$ eV) set to be greater than the Ga-Ga interaction energy ($E_{\text{Ga-Ga}} = 0.17$ eV), and model II does not include the surface exchange mechanism with $E_{\text{Al-Ga}} = E_{\text{Ga-Ga}} = 0.17$ eV. In the present study we use both models to perform growth simulations in which the V/III flux ratio is held constant. MC simulations were performed employing two different procedures (hereafter referred to as cases 1 and 2); in case 1 the opening of the Al shutter is accompanied by an instantaneous increase in the As flux to maintain a constant V/III flux ratio, and in case 2 the As flux is maintained constant and the Al and Ga fluxes are adjusted such that the V/III ratio remains constant at a given x_{Al} .

In order to match the conditions used in a recent experiment,¹⁰ MC simulations were performed for growth temperatures in the range of 950–990 K at a V/III flux ratio of 5. Growth simulations at a given temperature were initiated by first growing several layers of GaAs (on a As terminated GaAs surface) with incident Ga flux at 1 ML/s. The final configuration of the GaAs surface was stored and was subsequently used as the starting surface in the AlGaAs growth simulations. During the initial stages of GaAs

growth, the Ga desorption rate increased sharply and reached a steady state value when the Ga incorporation into the substrate was about 1 ML. For convenience, time zero was defined to be the instant at which the Ga desorption ratio attained steady state value for that growth temperature.

The AlGaAs growth simulations for the two cases described above were performed under the following conditions. In case 1 the incident flux for Ga was 1.0 ML/s and that for Al 0.667 ML/s; in case 2 the incident flux for Ga was 0.6 ML/s and that for Al 0.4 ML/s. Both cases yield a nominal composition of $\text{Al}_{0.4}\text{Ga}_{0.6}\text{As}$. In each simulation the Ga desorption ratio, defined as the fraction of incident Ga flux that desorbs, and the surface Al, Ga, and As fractions, defined as the fraction of surface sites terminated by the respective atomic species, were monitored as a function of time. The deposition of Ga was governed by the physisorption rule, while Al, which has a negligible desorption rate under the conditions of this study, was allowed enough jumps to find an appropriate SOS site. The Al-Ga surface exchange was implemented by having Al atoms that arrive at Ga terminated sites (by deposition or diffusion) replace the Ga atom. The ejected Ga atom then finds another site or desorbs, as per the physisorption rule.

III. RESULTS AND DISCUSSION

For completeness, we begin with a brief examination of results from growth simulations in which V/III flux ratio is not held constant. Figure 1(a) shows the time evolution of the Ga desorption ratio as predicted by models I and II. The experimental data are included for comparison. It is clear that a good match with experiment is obtained from both models. The origin of the transient in the Ga desorption ratio can be understood by examining the variation of the surface Al, Ga, and As concentrations with time, as shown in Fig. 1(b). The results from both models show that opening the Al shutter introduces an abrupt decrease in the As coverage due to a reduction in the V/III flux ratio. The observed transients can then be explained in terms of an abrupt increase in surface population of weakly bound Ga species followed by a decrease in the number of surface Ga sites, as described earlier.

Figure 2 shows the results from models I and II with the V/III flux ratios maintained constant using the two strategies specified in Sec. II. It is clear that both strategies eliminate the upward spike in the Ga desorption rate and result in a more rapid drop to the steady state value. However, both models predict that case 1 yields a steeper decline and a more steplike response in the Ga desorption behavior. The observed differences may be explained by examining the time evolution of the surface stoichiometry during growth. Figure 3 is a plot of the Al, Ga, and As concentrations of the surface versus time as predicted by model II. A comparison of the As coverage with that in Fig. 1(b) shows that both strategies are quite effective in eliminating drastic changes in the As coverage when the Al shutter is opened. Figure 3 also reveals that in case 2 the Ga surface concentration decreases more slowly and is also at a higher value than in case 1 before both profiles converge to the same steady state value.

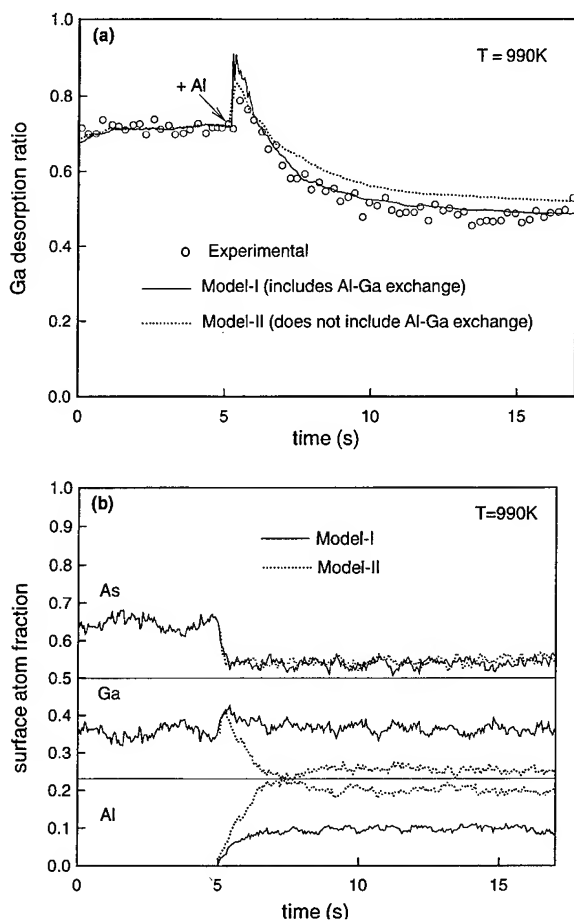


FIG. 1. Models result from the simulation with the V/III flux ratio not held constant showing the time evolution of the (a) Ga desorption ratio and (b) surface stoichiometry. Experimental data in (a) are from Ref. 10. The horizontal lines in (b) were included for ease in delineating profiles corresponding to Al, Ga, and As.

The predicted Ga surface concentration profiles are thus consistent with the Ga desorption profiles shown in Figs. 2(a) and 2(b). The slower decrease in the Ga surface concentration for case 2 may be explained as follows. Given that the V/III flux ratio is the same for both cases, and that the group-III flux during AlGaAs growth is higher for case 1 (1.667 ML/s vs 1.0 ML/s for case 2), it follows that the amount of excess As, which is simply the As flux minus the group-III flux, is higher for case 1. This greater excess of As causes an increased incorporation rate for surface Ga, resulting in the effect observed.

In order to examine the effectiveness of the two strategies employed, it would be useful to compare the stoichiometry of the AlGaAs-on-GaAs interface predicted by the different models. The Al concentration profiles predicted by models I and II are shown in Figs. 4(a) and 4(b), respectively. For the case in which the V/III ratio is not held constant, we observe significant buildup of the Al concentration on the GaAs surface. In model II this buildup is confined mostly to the top GaAs layer, and is caused by Al atoms occupying cation sites left vacant by desorbing Ga atoms. The buildup in Al concentration is even more pronounced in model I due to the

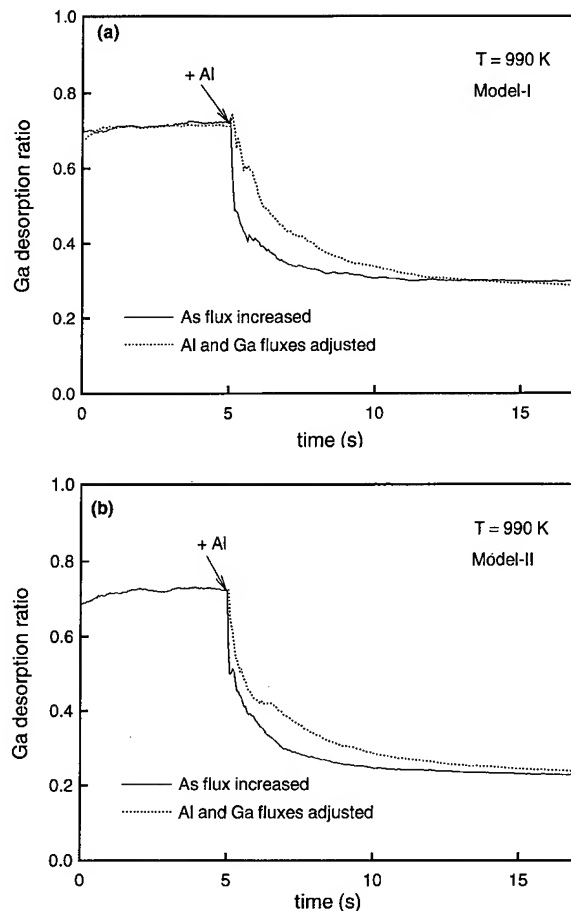


FIG. 2. Model results for the two growth strategies that maintain a constant V/III flux ratio, as predicted by (a) model I (includes Al-Ga surface exchange) and (b) model II (excludes Al-Ga exchange).

presence of the Al-Ga exchange mechanism. Both models predict that maintaining a constant V/III flux ratio results in an interface that is closer to the ideal AlGaAs-on-GaAs interface. The best results are achieved in case 1, however,

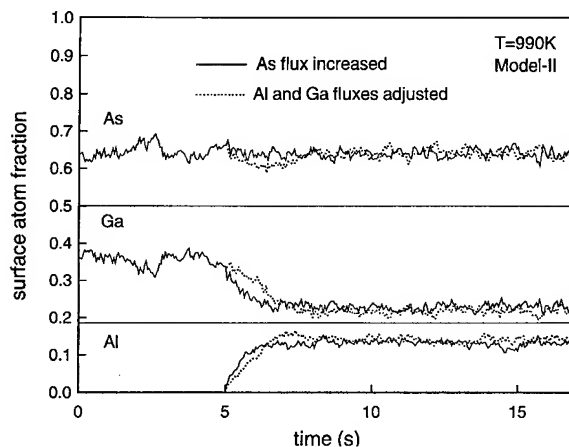


FIG. 3. Evolution of the surface stoichiometry as predicted by model II (excludes Al-Ga surface exchange) with the V/III flux ratio maintained constant by two different methods. The horizontal lines were included for ease in delineating profiles corresponding to Al, Ga, and As.

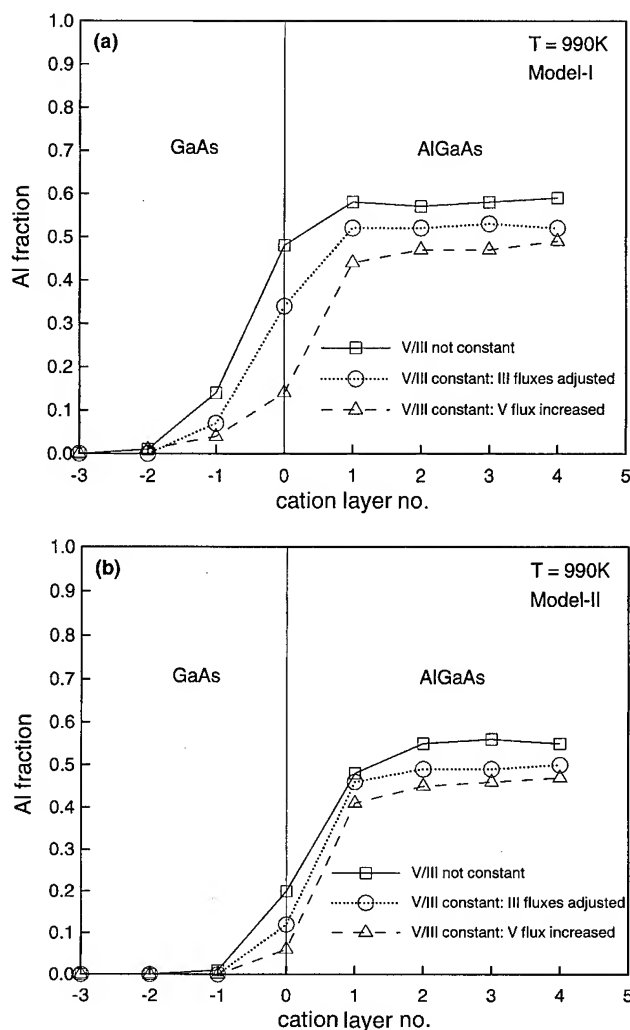


FIG. 4. Al concentration profiles for conditions with and without the V/III flux ratio maintained constant, as predicted by (a) model I (includes Al-Ga exchange) and (b) model II (excludes Al-Ga exchange). The zero of the abscissa denotes the initial GaAs surface layer over which AlGaAs growth is simulated. The vertical line denotes the ideal AlGaAs-on-GaAs interface.

indicating that increasing the As flux is more efficient at suppressing both the Al-Ga exchange mechanism and Ga desorption from the GaAs surface. The influence of the As flux on the Al-Ga exchange mechanism as predicted by model I is in agreement with experimental studies^{14,15} that report a similar effect of the As overpressure on the exchange mechanism.

It is worth noting that there is a good correlation between the Al concentration profiles shown in Fig. 4 and their respective Ga desorption profiles presented in Figs. 1(a) and 2. The steeper decline in the Ga desorption ratio leads to a more abrupt interface. It should be noted, however, that a sharp rise in the Ga desorption rate (for conditions in which the V/III flux ratio is not held constant) does not lead to an overenrichment of Al at the interface as was predicted in a previous model.⁹ The previous model does not take into con-

sideration the penetration of impinging Al atoms into the GaAs surface and thus overestimates the Al enrichment at the AlGaAs-on-GaAs interface. Our models suggest that the spike in the Ga desorption profile leads to the formation of a diffuse interface of the order of 2–3 ML. It should be pointed out that the predictions of different models can be verified by performing appropriate growth experiments. While *in situ* DMS can be used to verify the predicted Ga desorption profiles, the stoichiometric profiles can be verified by high-resolution transmission electron microscopy using the chemical lattice imaging technique.¹⁶

IV. SUMMARY

The transient in the Ga desorption rate during AlGaAs-on-GaAs heterointerface formation is caused by a reduction in the V/III flux ratio that coincides with the opening of the Al shutter. The present MC study shows that “steplike” variations in the Ga desorption rate and abrupt interfaces can be realized by adopting conditions in which the V/III ratio is maintained constant during growth. The best results are obtained when the reduction in the V/III ratio is compensated by an increase in the As flux.

ACKNOWLEDGMENTS

The authors are grateful to Professor Stewart Harris for many useful discussions. The research was supported by Air Force Contract No. F33615-94-C-5804 under the sponsorship of the Materials Directorate, Wright Laboratory, WL/MLPO, at Wright-Patterson Air Force Base, Ohio.

¹W. T. Tsang, *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1985), Vol. 1-22, Part A, p. 95.

²H. Morkoç, T. J. Drummond, W. Kopp, and R. Fischer, *J. Electrochem. Soc.* **129**, 824 (1982).

³C. T. Foxon, *Heterojunctions and Semiconductor Superlattices*, edited by G. Allen, G. Bastard, N. Boccard, M. Lannoo, and M. Voos (Springer, Berlin, 1986), p. 27.

⁴A. H. Kean, C. R. Stanely, M. C. Holland, J. L. Martin, and J. N. Chapman, *J. Cryst. Growth* **111**, 189 (1991).

⁵J. P. Reithmaier, R. F. Broon, and H. P. Meier, *Appl. Phys. Lett.* **61**, 1222 (1992).

⁶E. M. Gibson, C. T. Foxon, J. Zhang, and B. A. Joyce, *Appl. Phys. Lett.* **57**, 1203 (1990).

⁷J. Zhang, E. M. Gibson, C. T. Foxon, and B. A. Joyce, *J. Cryst. Growth* **111**, 93 (1991).

⁸K. R. Evans, C. E. Stutz, D. K. Lorance, and R. L. Jones, *J. Vac. Sci. Technol. B* **7**, 259 (1989).

⁹K. R. Evans, C. E. Stutz, E. N. Taylor, and J. E. Ehret, *J. Vac. Sci. Technol. B* **7**, 2428 (1991).

¹⁰K. Mahalingam, D. L. Dorsey, K. R. Evans, and R. Venkatasubramanian, *J. Cryst. Growth* (in press).

¹¹A. Madhukar and S. V. Ghaisas, *CRC Critical Reviews of Solid State and Materials Sciences* (Chemical Rubber, Cleveland, 1988), Vol. 14, p. 1.

¹²J. Singh and K. K. Bajaj, *J. Vac. Sci. Technol. B* **3**, 520 (1985).

¹³W. A. Harrison, *Electronic Structure and Properties of Solids* (Dover, New York, 1989), p. 176.

¹⁴T. Saitoh, M. Tamura, and J. E. Palmer, *Inst. Phys. Conf. Ser.* **141**, 345 (1995).

¹⁵T. Saitoh, A. Hashimoto, and M. Tamura, *Appl. Surf. Sci.* **60/61**, 228 (1992).

¹⁶A. Ourmazd, F. H. Baumann, M. Bode, and Y. Kim, *Ultramicroscopy* **34**, 237 (1990).

Kink defects and Fermi level pinning on (2×4) reconstructed molecular beam epitaxially grown surfaces of GaAs and InP studied by ultrahigh-vacuum scanning tunneling microscopy and x-ray photoelectron spectroscopy

Yasuhiko Ishikawa,^{a)} Takashi Fukui, and Hideki Hasegawa

Research Center for Interface Quantum Electronics and Graduate School of Electronics and Information Engineering, Hokkaido University, N13 W8, Sapporo 060, Japan

(Received 13 January 1997; accepted 3 April 1997)

The relationship between kink defects and Fermi level pinning on molecular beam epitaxially grown GaAs and InP(001)- (2×4) surfaces is studied in detail by scanning tunneling microscopy and x-ray photoelectron spectroscopy. In Si-doped GaAs, the kink density increased with doping as previously found. However, actual density depended very much on the reconstruction phases and experimental conditions. At high Si doping levels, Fermi level was strongly pinned below mid-gap, but the measured kink density was found not to be large enough to explain pinning by the previous kink-deep-acceptor model assuming that each kink forms a single discrete level. In Si-doped InP, the kink density remained constant with the increase of Si doping, although the Fermi level was pinned above mid-gap. The result cannot be explained by the kink-deep-acceptor model either.

© 1997 American Vacuum Society. [S0734-211X(97)09604-2]

I. INTRODUCTION

In III-V compound semiconductor based microelectronics, control of the electronic properties of surfaces and interfaces is one of the most important issues for further progress. Namely, III-V compound semiconductor surfaces normally possess high density of surface states which pin the position of the surface Fermi level within a certain narrow energy range. This Fermi level pinning phenomenon makes the control of the Schottky barrier heights by metal work function difficult as well as prevents formation of well-behaved insulator-semiconductor interfaces.

Because of this importance, a number of studies on the surface states and the Fermi level pinning have been carried out previously, and several models such as unified defect model (UDM),¹ metal-induced gap state (MIGS) model,^{2,3} disorder-induced gap state (DIGS) model,⁴ and effective work function (EWF) model⁵ have been proposed concerning their origin and the pinning mechanism. However, these models lacked direct microscopic justification. In this respect, Pashley *et al.*^{6,7} have recently found by scanning tunneling microscope (STM) observation that the density of the so-called kink defects in the missing-dimer arrangements on *n*-type GaAs(001)- (2×4) surface grown by molecular beam epitaxy (MBE) have a strong correlation with the Si doping. Based on this observation, these authors proposed that each kink acts as a single discrete deep surface acceptor site and a high density of kinks cause mid-gap Fermi level pinning on heavily Si-doped *n*-type GaAs(001)- (2×4) surfaces. Later, a scanning tunneling spectroscopy (STS) study⁸ detected directly the existence of negative charge near kinks, further confirming the acceptor nature of kinks. This model is hereafter referred to as the kink-deep-acceptor model. This model

seems to be the first microscopic model concerning the Fermi level pinning. However, the quantitative validity of the model as well as its applicability to other III-V compound semiconductor surfaces has not been well established so far. Only kinks on surfaces of narrow gap InAs have been investigated by Yamaguchi and Horikoshi,⁹ where basically the same kink-deep-acceptor model was used with certain modifications to take account of the narrow gap nature of the material.

The purpose of this article is to study the relationship between the kink defect density and the Fermi level pinning on GaAs and InP (2×4) surfaces prepared by MBE in detail by using STM and x-ray photoelectron spectroscopy (XPS), and to investigate the quantitative validity and applicability of the kink-deep-acceptor model.

It is shown here that the kink density in Si-doped GaAs increases with doping as previously found. However, actual density was found to depend very much on the reconstruction phases and experimental conditions. At high Si doping levels, Fermi level is strongly pinned below mid-gap, but the measured kink density was found not large enough to explain pinning in terms of the kink-deep-acceptor model.

In Si-doped InP, the kink density was low and remained constant with the increase of Si doping, although the Fermi level was found to be pinned above mid-gap. The result cannot be explained by the kink-deep-acceptor model either.

II. EXPERIMENT

A. Experimental system and MBE growth of samples

An ultrahigh vacuum (UHV)-based multi-chamber system was used. In this system, a solid source MBE chamber, a gas source (GS) MBE chamber, XPS (Perkin-Elmer PHI 5100C) chamber, and UHV-STM (JEOL JSTM-4600) cham-

^{a)}Electronic mail: ishikawa@ryouko.rciqe.hokudai.ac.jp

ber are connected by a UHV transfer chamber, together with other chambers. Details of this system have been described elsewhere.^{10,11}

For preparation of GaAs surfaces, surface oxides were desorbed from n^+ ($>1 \times 10^{18} \text{ cm}^{-3}$)-GaAs (001) substrates in an As_4 flux at 600 °C first, and then thick Si-doped GaAs epitaxial layers ($>5000 \text{ \AA}$) were grown in a conventional solid source MBE. According to Ref. 6, this thickness should be large enough for saturation of the kink density. The growth rate was 1.0 ML/s, and the growth temperature was 580 °C. Reconstruction patterns were monitored by reflection high-energy electron diffraction (RHEED) with an electron energy of 15 keV. RHEED patterns were streaky (2×4) during the growth.

In addition to the Si-doped GaAs samples grown under the above condition, three different types of GaAs samples were also prepared in order to compare them with Si-doped samples. They were a Be-doped p -type sample grown at 580 °C, a Si and Be co-doped n -type sample grown at 580 °C, and a Si-doped n -type sample grown at a low temperature (LT) of 300 °C, respectively.

For preparation of InP surfaces, the surface oxides were desorbed from n^+ -InP (001) substrates in a phosphorus flux at 500 °C, and then thick Si-doped n -type InP epitaxial layers were grown in a GSMBE using tertiarybutylphosphine (TBP), metallic In, and elemental Si at 470 °C with a growth rate of 0.5 ML/s. TBP was precracked at 750 °C. Clean p -type InP (2×4) surfaces were also prepared on Zn-doped p -type substrates by thermal desorption of surface oxides in a phosphorus flux at 500 °C.

Doping concentration was calibrated by the Hall effect measurements for both GaAs and InP materials. For Si-doped GaAs, the observed electron concentration was almost equal to the Si doping concentration up to a Si doping of $(5-7) \times 10^{18} \text{ cm}^{-3}$ beyond which the electron concentration saturated in agreement with a previous report.¹² On the other hand, for Be-doped p -type GaAs and Si-doped n -type InP, the carrier concentration did not show saturation even if the doping was done above the $(5-7) \times 10^{18} \text{ cm}^{-3}$ level. This also agrees with the previous reports.^{13,14}

After the growth was completed, the samples were annealed in an arsenic or a phosphorus flux for 3–10 min before cooling. STM observations were done on samples cooled down to room temperature.

B. Preparation of various GaAs(001)-(2×4) surfaces for room temperature observation

(2×4) GaAs surfaces for room temperature STM observation were prepared by the following two procedures. The first procedure was to establish the initial reconstruction phase by annealing the surface after MBE growth at a suitable temperature and then to cool the sample by gradually reducing the intensity of the As_4 flux irradiating the surface during cooling in such a way that the initial RHEED pattern was maintained during cooling down to room temperature. In this study, three (2×4) phases, i.e., α , β , and γ phases of GaAs were realized. They could be clearly distinguished by

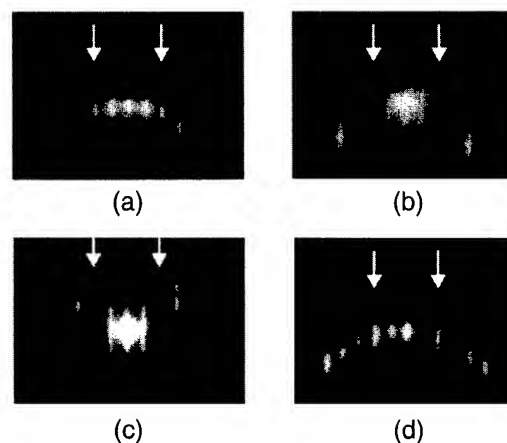


FIG. 1. Typical $[110]$ RHEED patterns for (a) GaAs (2×4) β phase, (b) GaAs (2×4) γ phase, (c) heavily Si-doped ($1.0 \times 10^{19} \text{ cm}^{-3}$) GaAs (2×4) surface obtained under the same condition as the β phase, and (d) InP (2×4) surface.

the intensity of the $2/4$ fractional order streaks in the $[110]$ RHEED pattern.¹⁵ Namely, an α phase with the weak $2/4$ fractional order streaks in the $[110]$ RHEED pattern appeared above 610 °C, a β phase with the strong $2/4$ fractional order streaks appeared at around 580 °C, and a γ phase with the weak $2/4$ fractional order streaks appeared at around 520 °C, respectively. All these patterns were maintained down to room temperature by proper adjustment of As_4 flux intensity during cooling.

The second procedure was to cool the sample without reducing the As_4 flux down to room temperature where the (2×4) pattern changed to the $c(4 \times 4)$ pattern at about 500 °C. After the background As was sufficiently removed, the sample was heated up to 420 °C in the absence of an arsenic flux where $c(4 \times 4)$ reconstruction changed to (2×4) γ reconstruction. This γ -phase pattern could be maintained down to room temperature by subsequent cooling. In this study, the kink densities on both β and γ phase (2×4) surfaces were studied in detail, changing the Si doping levels in the epitaxial layers over a wide range.

Figures 1(a) and 1(b) show the typical $[110]$ RHEED patterns of GaAs (2×4) β and γ phases obtained at room temperature on undoped or moderately Si-doped samples, respectively. The pattern from the β phase shown in Fig. 1(a) possesses a clear Laue circle. On the other hand, the pattern from the γ phase does not possess a clear Laue circle. This nature of the γ -phase pattern was independent of the above two procedures.

However, in heavily Si-doped (above around $5 \times 10^{18} \text{ cm}^{-3}$) samples, such clear β -phase pattern with a Laue circle as in Fig. 1(a) could not be obtained under the same preparation procedure. Instead, a pattern shown in Fig. 1(c) was obtained where only very weak $2/4$ fractional order streaks appeared without Laue circle. Appearance of a similar pattern at high Si doping was reported in a previous study.¹⁶ Although the pattern in Fig. 1(c) is different from that in Fig. 1(a) and more like a γ -phase pattern with sharpened streaks, we refer to this pattern as β^* phase in this

article for simplicity to indicate the preparation condition. The transition from this β^* phase to γ phase in heavily Si-doped samples could be clearly distinguished by broadening the streaks in the $[\bar{1}10]$ RHEED pattern.

C. Preparation of InP(001)-(2×4) surfaces at room temperature

On InP (2×4) surfaces, the $[\bar{1}10]$ RHEED pattern at high temperature always showed β -phase patterns with high intense 2/4 fractional order streaks. This (2×4) surface for room temperature observation was obtained in the following way. After the growth, the surface was cooled without changing the phosphorus flux intensity down to 370 °C where the P-excess (2×1) surface appeared. Then, the phosphorus flux was stopped, and the sample was annealed at 370 °C in the MBE chamber until residual phosphorus vapor was sufficiently removed, and the (2×4) pattern was recovered.^{10,11} Then the sample was cooled down to room temperature where the β phase (2×4) surfaces with strong 2/4 fractional order streaks and a clear Laue circle in the $[\bar{1}10]$ RHEED pattern was obtained as shown in Fig. 1(d). Even in the case of the heavily Si-doped samples, InP surfaces showed well-defined β -phase patterns with strong 2/4 fractional streaks.

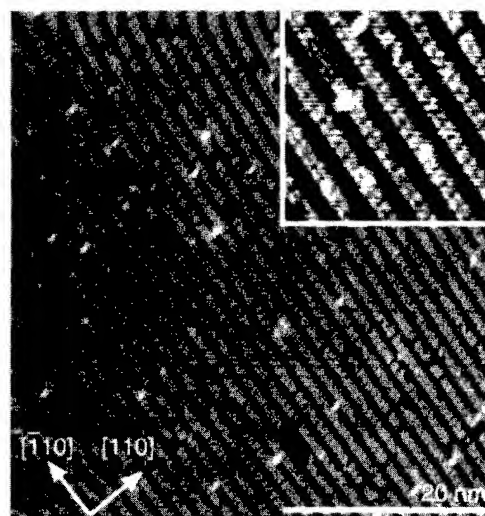
D. Conditions of STM and XPS measurements

After the desired (2×4) RHEED pattern was obtained at room temperature, the samples were rapidly transferred to STM or to XPS chambers through the UHV transfer chamber. In this study, a sample chip with an area of 5 mm×5 mm for the STM measurement and another chip with an area of 10 mm×10 mm for the XPS measurement were placed next to each other on the same holder in the MBE chamber and underwent exactly the same growth and surface preparation procedures down to room temperature before they were finally separated into different chambers. For this purpose, a special sample holder which was separable in the UHV environment was used.

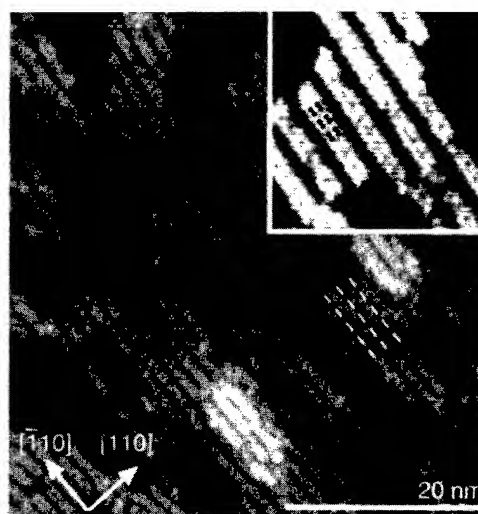
Room temperature STM observations were made using electrochemically polished tungsten tips with a constant current mode (0.10–0.20 nA at sample bias of –2.5 to –2.0 V, i.e., filled state images). Our STM equipment had high lateral and vertical resolutions enough to discuss the atomic arrangements of III–V compound surfaces as previously described.¹¹

For the kink density measurements, 5–10 measurement spots were chosen on the STM sample chip, and eight images were taken on each spot by shifting the probe position. The average kink density and its rms deviation were determined from statistical analysis of 40–80 images.

XPS measurements were done at room temperature using Al $K\alpha$ radiation ($h\nu=1486.6\text{eV}$) for GaAs, and Mg $K\alpha$ radiation ($h\nu=1253.6\text{eV}$) for InP, in order to measure the surface Fermi level position. The binding energies of the spectra were carefully calibrated by the separate measurements of Au 4f_{7/2} peak position.



(a)



(b)

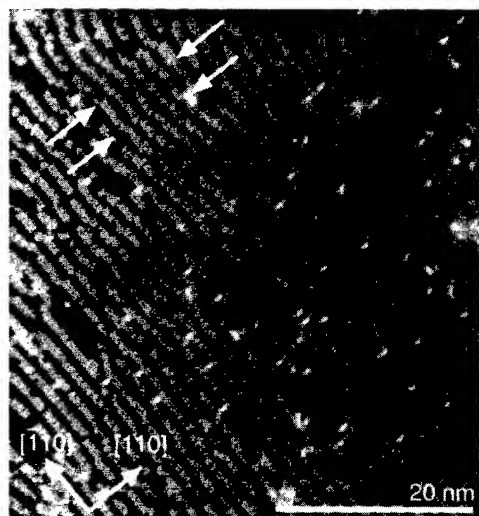
FIG. 2. STM images of Si-doped GaAs (2×4) surfaces for (a) β phase and (b) γ phase. Si concentrations of the samples were (a) $1.0 \times 10^{18} \text{ cm}^{-3}$ and (b) $3.0 \times 10^{17} \text{ cm}^{-3}$. The insets show parts of the both images under higher magnification.

III. EXPERIMENTAL RESULTS

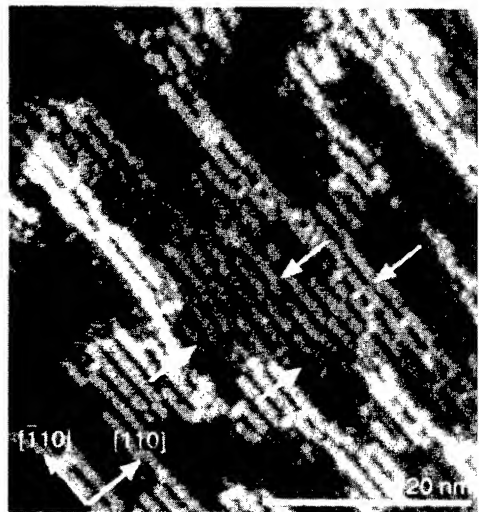
A. Kink density on GaAs(001)-(2×4) surfaces

Typical STM images taken on the (2×4) surfaces of highly Si-doped GaAs corresponding to β and γ phases are shown in Figs. 2(a) and 2(b), respectively. The Si doping concentrations were $1.0 \times 10^{18} \text{ cm}^{-3}$ for Fig. 2(a) and $3.0 \times 10^{17} \text{ cm}^{-3}$ for Fig. 2(b), respectively. The insets show parts of both images under higher magnifications.

The unit cells of both surfaces possess rows of two As dimers and rows of two missing As dimers in the top layer, as previously reported by Avery *et al.*¹⁷ It was found that the β -phase surface possesses very large terraces where the straight dimer rows and missing-dimer rows, seen as white



(a)



(b)

FIG. 3. STM images of heavily Si-doped ($1.0 \times 10^{19} \text{ cm}^{-3}$) GaAs (2 \times 4) surfaces for (a) β^* phase and (b) γ phase.

and dark stripes on the STM images in Fig. 2(a), respectively, extended over long distances along the $[110]$ direction. Thus, the density values of kinks on the β -phase surfaces were low. On the other hand, high density of kinks as well as many holes and islands were observed on γ -phase surfaces.

By further increasing the Si doping level, STM images shown in Figs. 3(a) and 3(b) were obtained on β^* - and γ -phase surfaces, respectively. The Si doping concentration was $1.0 \times 10^{19} \text{ cm}^{-3}$ for both samples. A large number of kinks are observed on both images. As seen in Figs. 3(a) and 3(b), there is a strong tendency on both images for a series of kinks to be formed next to each other, aligning approximately along the $[110]$ direction as shown by the arrows. At each kink site, both of the dimer rows and missing-dimer rows are shifted horizontally by one atomic spacing.

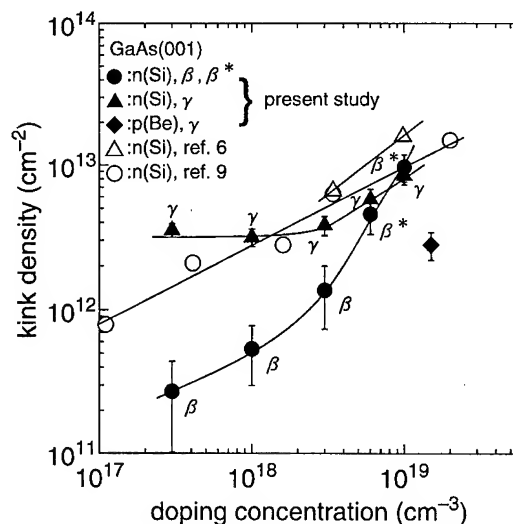


FIG. 4. Kink density vs doping concentration of GaAs (2 \times 4) β -, β^* - and γ -phase surfaces. The data obtained by Pashley and Haberern (see Ref. 6) (Δ) and Yamaguchi and Horikoshi (Ref. 9) (\circ) were included.

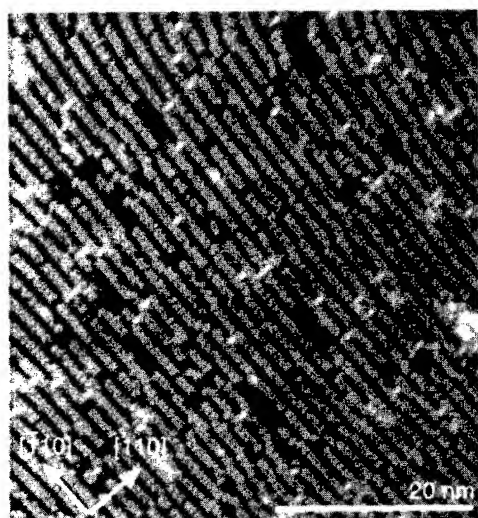
The average kink densities obtained from GaAs (2 \times 4) images are summarized in Fig. 4 for n -type surfaces with various Si doping concentrations and for a Be-doped p -type sample with γ phase. As mentioned previously, each data point represents an average of 40–80 STM images, and the error bars show the rms deviation values of the measured density data. The experimental results by other workers^{6,9} are also included. Judging from the STM images, it seems that β - and β^* -phase surfaces were used by these other two groups, although it is not explicitly stated in the papers.

As seen in Fig. 4, the kink density observed on Si-doped GaAs (2 \times 4) β -, β^* -, and γ -phase surfaces increases with Si doping concentration monotonously in qualitative agreement with the observations of other groups. However, the values of the kink density and the slope of the curves are quite different. It is also seen that the behavior of the kink density is strongly dependent on the reconstruction phases. Namely, on Si-doped γ -phase surfaces, the kink density increases with Si doping concentration above $\sim 3 \times 10^{18} \text{ cm}^{-3}$, but remains constant at about $3 \times 10^{12} \text{ cm}^{-2}$ for Si doping below $3 \times 10^{18} \text{ cm}^{-3}$. This constant value is much larger than those on the β -phase surfaces.

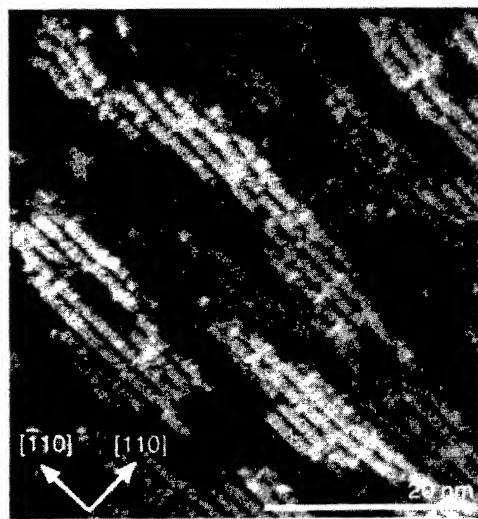
The Be-doped p -type γ -phase surface with the Be concentration of $1.5 \times 10^{19} \text{ cm}^{-3}$ showed unexpectedly high kink density of about $3 \times 10^{12} \text{ cm}^{-2}$, although this value is much lower than those on Si-doped n -type samples with the same level of doping.

B. Kink density on Si-doped LT-grown GaAs and Si and Be co-doped GaAs(001)-(2 \times 4) surfaces

Figures 5(a) and 5(b) show the STM images of a Si-doped LT-grown GaAs (2 \times 4) β^* -phase surface and a Si and Be co-doped GaAs (2 \times 4) γ -phase surface, respectively. The (2 \times 4) surface on an LT-grown sample was formed by annealing the surface in an As_4 flux at 580 $^\circ\text{C}$ for 10 min after



(a)



(b)

FIG. 5. STM image of GaAs (2×4) surfaces for (a) Si-doped β^* -phase surface grown at low temperature (300 °C) and (b) Si and Be co-doped γ -phase surface grown at 580 °C. Si doping concentrations are $6.0 \times 10^{18} \text{ cm}^{-3}$ for both samples.

the LT growth. The Si doping concentration was the same and equal to $6.0 \times 10^{18} \text{ cm}^{-3}$ for both of the LT and co-doped samples, but the electron concentrations were found to be different, being $2.0 \times 10^{18} \text{ cm}^{-3}$ for the LT-grown GaAs and $1.4 \times 10^{17} \text{ cm}^{-3}$ for the Si and Be co-doped sample, respectively.

High densities of kink defects were observed on both of these surfaces in spite of reduced electron concentrations. The average kink densities were $4.3 \times 10^{12} \text{ cm}^{-2}$ for the LT-grown GaAs and $5.4 \times 10^{12} \text{ cm}^{-2}$ for the co-doped GaAs, respectively. The kink densities on these surfaces are compared in Fig. 6 with those of the sample with only Si doping to the same level of $6.0 \times 10^{18} \text{ cm}^{-3}$ as a function of the net electron concentration. The solid line is the theoretical line

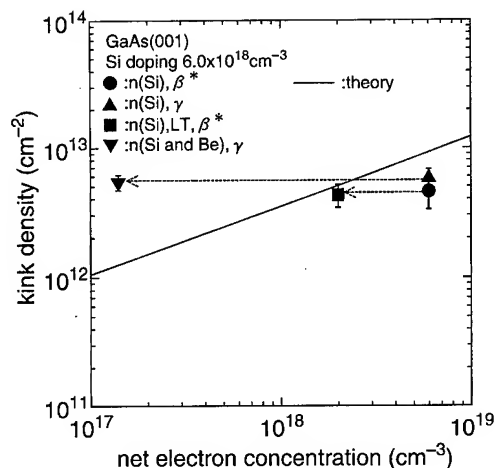


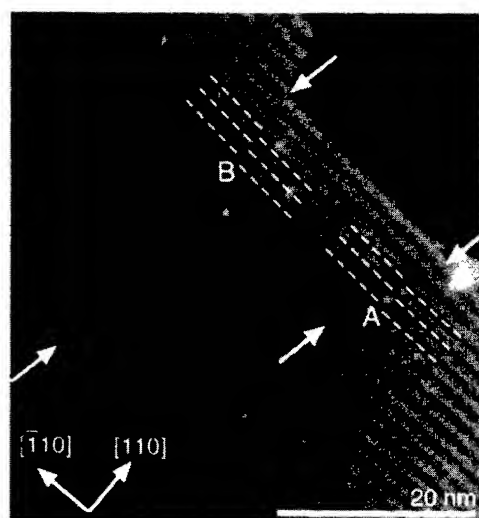
FIG. 6. Kink density vs net carrier concentration for GaAs (2×4) surfaces. All samples were doped with Si of $6.0 \times 10^{18} \text{ cm}^{-3}$. The solid line is the theoretical line explained later.

explained later. As seen in Fig. 6, the kink densities on LT-grown GaAs and Si and Be co-doped GaAs samples were found to be almost equal to those obtained under the conventional Si doping only. This result indicates that the kink density depends principally on the Si doping concentration itself under the same reconstruction phase, and does not depend on the net electron concentration.

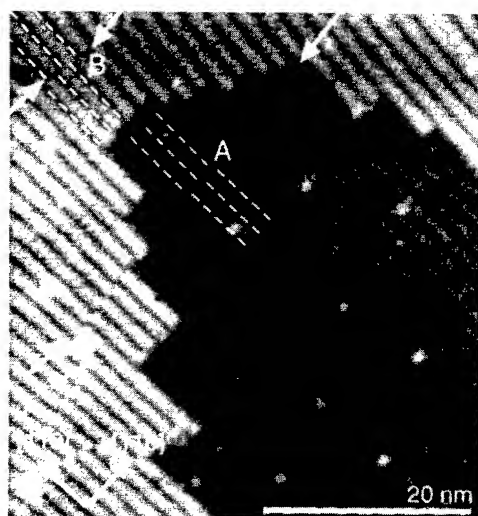
C. Kink density on InP(001)-(2×4) surfaces

Figures 7(a) and 7(b) show two STM images of Si-doped InP (2×4) surfaces with the doping concentration of 4.5×10^{18} and $1.4 \times 10^{19} \text{ cm}^{-3}$, respectively. As seen in Figs. 7(a) and 7(b), the missing-dimer rows on InP surfaces are very regular even at high doping levels. They extend over long distances straightly along [110] direction, being similar to the case of the lightly Si-doped GaAs (2×4) β -phase surface shown in Fig. 2(a). Kinks in missing-dimer rows were observed also on these surfaces. But, their densities are low. They tend to form next to each other as in the case of GaAs (2×4) surfaces. However, the amount of horizontal shifts are not only one atomic spacing as shown by the letter A as in the case of GaAs, but also can be two atomic spacings, as shown by the letter B in the images. This indicates that the atomic arrangements around kink sites on the InP (2×4) surface may be partially different from those on the GaAs (2×4) surface.

Figure 8 shows the measured average kink densities on InP (2×4) surfaces as a function of the Si doping concentration with error bars showing rms deviations. Here the number of all of the kink sites were included, being irrespective of type A and B shifts. As seen in Fig. 8, kinks in missing-dimer rows showed no correlation with the Si doping concentration in the case of InP, giving a low constant value of around $1 \times 10^{12} \text{ cm}^{-2}$. In addition, the Zn-doped p-InP (2×4) surface also showed nearly the same kink density.



(a)



(b)

FIG. 7. STM images of Si-doped InP (2×4) surfaces. Si doping concentrations are (a) $4.5 \times 10^{18} \text{ cm}^{-3}$ and (b) $1.4 \times 10^{19} \text{ cm}^{-3}$, respectively.

D. Fermi level positions at GaAs and InP (2×4) surfaces

The surface Fermi level positions of the GaAs surfaces were determined from the XPS valence band spectra. The surface photovoltage effect,¹⁸ which is known to reduce the band bending during measurement, was ignored. According to calculation,¹⁸ such an effect appears to be negligible at room temperature particularly for surfaces with heavily *n*-type doping with high surface recombination. The Fermi level positions of the surfaces from the valence band maximum (VBM) measured on the Si-doped *n*-type GaAs (2×4) surfaces are summarized in Fig. 9(a) as a function of Si doping concentration.

As seen in Fig. 9(a), the surface Fermi level positions lay around the mid-gap for the lightly Si-doped ($1.0 \times 10^{17} \text{ cm}^{-3}$) GaAs (2×4) surfaces, and a slightly deeper position

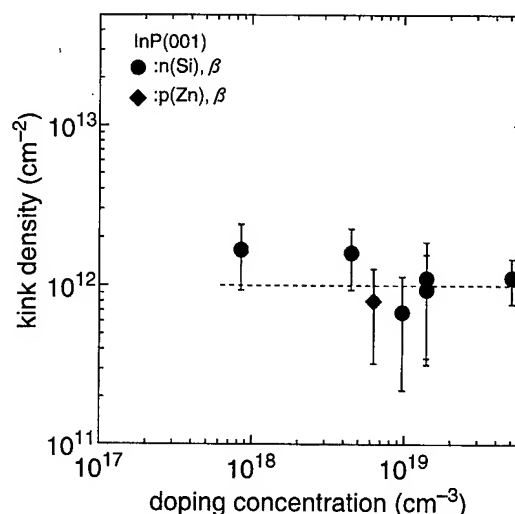


FIG. 8. Kink density vs doping concentration of InP (2×4) surfaces.

of about $E_V + 0.62 \text{ eV}$ below the mid-gap as the Si doping level is increased. The result seems to be independent of reconstruction phases. Figure 9(b) shows the measured Fermi level positions on InP (2×4) surfaces as a function of the doping concentration. The surface Fermi level positions were determined from $P2p_{3/2}$ core level peak positions using the relation of $E_{P2p_{3/2}} - E_V = 127.74 \text{ eV}$.¹⁹ The apparent shifts of the peak energy positions caused by the band bending within the electron escape depth²⁰ were corrected by numerical analysis. Details of this correction procedure were described previously.¹¹ It is seen from Fig. 9(b) that the Fermi level on the *n*-InP (2×4) surface lies at $E_F = E_V + 0.95 \text{ eV}$, and that on *p*-InP the (2×4) surface lies at $E_F = E_V + 0.75 \text{ eV}$ on average. These values are far from the flatband condition and indicate the presence of Fermi level pinning.

IV. DISCUSSION

A. Applicability of kink-deep-acceptor model for Fermi level pinning on GaAs (2×4) surface

Previous models on surface Fermi level pinning which assume the presence of surface states can be classified into two types. Namely, one type assumes the presence of discrete deep defect states at surface as in the case of the unified defect model (UDM) by Spicer *et al.*¹ The other type assumes U-shaped continuum of surface states with a charge neutrality level as in the case of the metal-induced gap state (MIGS) model³ and the disorder-induced gap state (DIGS) model.⁴

Let us consider *n*-type materials. Then, in the former case, pinning is caused by discrete deep acceptor states as schematically shown in Fig. 10(a). The condition for the surface Fermi level E_{FS} to be pinned in the close vicinity of the surface acceptor level E_{SA} is obviously

$$N_{SA} \geq N_{\text{depl}} \bigg|_{E_{FS} = E_{SA}}, \quad (1)$$

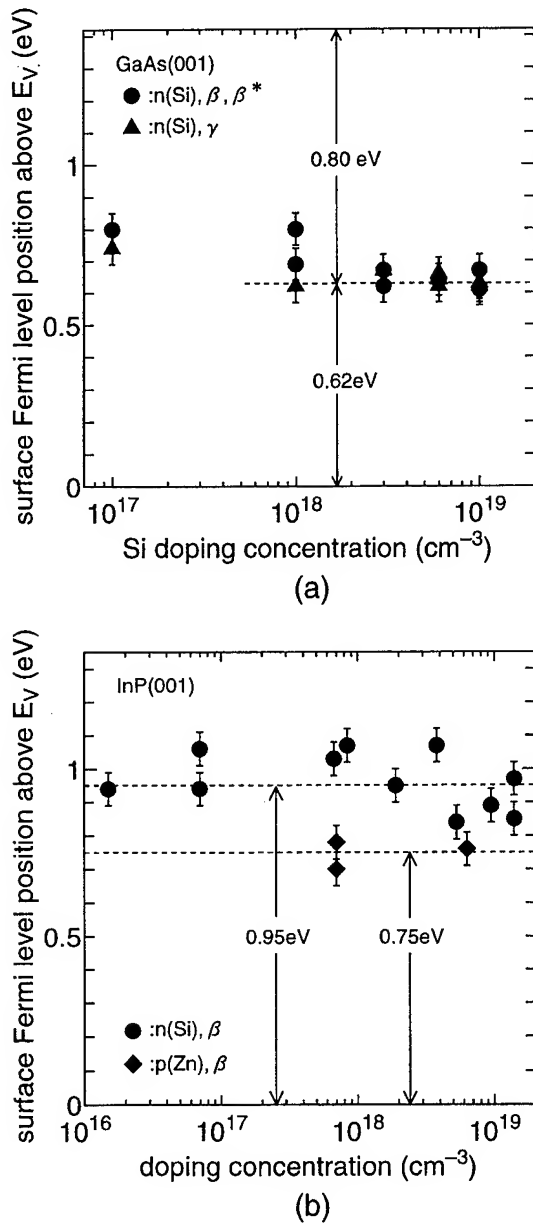


FIG. 9. Surface Fermi position above E_V measured by XPS as a function of doping concentration for (a) GaAs (2×4) surface and for (b) InP (2×4) surface.

where N_{SA} is the area density of surface acceptors and the $N_{depl}|_{E_{FS}=E_{SA}}$ is the total number per unit area of ionized donors when E_{FS} agrees with E_{SA} . Here, N_{depl} is given in the depletion approximation as a function of E_{FS} by

$$N_{depl}(E_{FS}) = \sqrt{2\epsilon\epsilon_0(E_C - E_{FS} - eV_n)N_d/e}, \quad (2)$$

where e is the electron charge, ϵ is the dielectric constant of the semiconductor, ϵ_0 is the permittivity in vacuum, E_C is the conduction band minimum (CBM), eV_n is the distance between the CBM and the bulk Fermi level, and N_d is the ionized donor concentration. When the condition (1) is not met, E_{FS} is not pinned at E_{SA} , but lies above E_{SA} .

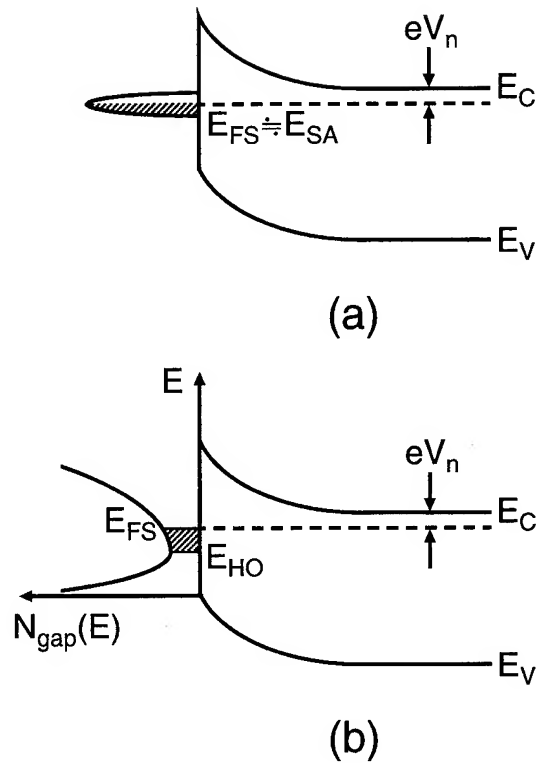


FIG. 10. Models of gap state distributions; (a) discrete distribution and (b) U-shaped continuous distribution.

On the other hand, in the case of gap state continuum, pinning takes place by the acceptor type gap state continuum as shown in Fig. 10(b) and the pinning position is determined by the following equation

$$\int_{E_{HO}}^{E_{FS}} N_{gap}(E) dE = N_{depl}(E_{FS}), \quad (3)$$

where $N_{gap}(E)$ is the surface state density per unit energy and area and E_{HO} is the charge neutrality level of the surface state continuum.⁴ In this continuum, surface state is of acceptor type above E_{HO} , and of donor type below E_{HO} , respectively. Note that Eq. (3) applies only to pinning at free surface and insulator-semiconductor interface, but not to pinning at metal-semiconductor interface where charge on metal has to be included.²¹

The kink-deep-acceptor model originally proposed by Pashley and Haberern⁶ to explain Fermi level pinning on n -type (2×4) GaAs surfaces belongs to the former type of pinning by discrete states where each kink acts as a discrete surface deep acceptor state which accommodates one electron per kink site. When the doping density is high, the high density of kinks pin the surface Fermi level at the energy position of the deep acceptor which is assumed to lie near mid-gap. In this model, there is a further assumption that each kink acceptor is formed so as to lower the energy of each electron from each Si donor by trapping the electron at a low lying acceptor level. This continues until E_{FS} agrees

with E_{SA} so that there is no energy gain. Thus, in this case, the number of the kink acceptor, N_{kink} , should satisfy

$$N_{\text{kink}} = N_{\text{depl}} \quad (4)$$

In order to compare the experimentally observed kink densities with the prediction of Eq. (4), the calculated values of N_{kink} using Eqs. (2) and (4) are shown for various pinning positions by the solid lines in Fig. 11(a). Data by other groups^{6,9} are also included. As seen in Fig. 11(a), there is a general trend that the kink density N_{kink} increases with the Si doping. This agrees with the observation by other workers.^{6,9} However, for a given doping, the observed kink density is very different among workers and among reconstruction phases. There is absolutely no indication that there exists one-to-one correspondence between the Si doping concentration and the kink density in spite of the presence of strong Fermi level pinning near mid-gap in all the samples.

Since Pashley *et al.*⁶⁻⁸ claimed the kinks become dominant pinning sites in the high Si doping range of 10^{18} – 10^{19} cm^{-3} , the surface Fermi level positions expected from the measured kink densities after the kink-deep-acceptor model and those directly measured by XPS are compared in detail in Fig. 11(b). This was done only for our data due to ambiguity in the exact Fermi level positions in previous works.^{6,9} It should be noted that, as mentioned previously, our samples for kink measurement and those for XPS measurement for Fermi level pinning were prepared in this study next to each other and underwent exactly the same growth and cooling procedure, so that possible change of the kink density during cooling should not affect the measurement. In Fig. 11(b), data pairs taken on the neighboring samples are connected by dashed lines. For the sake of easier understanding of the data, doping values are shifted horizontally within doping measurement accuracy.

If the kink-deep-acceptor model is correct, these two sets of E_{FS} values should coincide with the measurement accuracy. However, it is apparent from Fig. 11(b) that the two surface Fermi level positions are very different and differences are much larger than the measurement accuracy, particularly in the case of well-behaved β , and β^* surfaces. Thus, the present result cannot be explained quantitatively by the kink-deep-acceptor model.

The result shown in Fig. 6 concerning the relation between the net electron concentration and kink density is also very much against the kink-deep-acceptor model. This is because the reduction in the net ionized donor density in the surface depletion layer due to compensation should result in reduction of the surface kink acceptors. The theoretical relationship using Eqs. (2) and (4) with the assumption that the net electron concentration is equal to the net ionized donor concentration is shown by the solid line in Fig. 6, using the observed Fermi level position of $E_F = E_V + 0.62$ eV. However, no change of the kink density was observed, and it is as though the kink density is determined by the number of Si atoms itself, not by the number of net ionized impurity atoms. Of course, there is a possibility that the number of the compensating acceptors, i.e., Be in the Be-doped case and

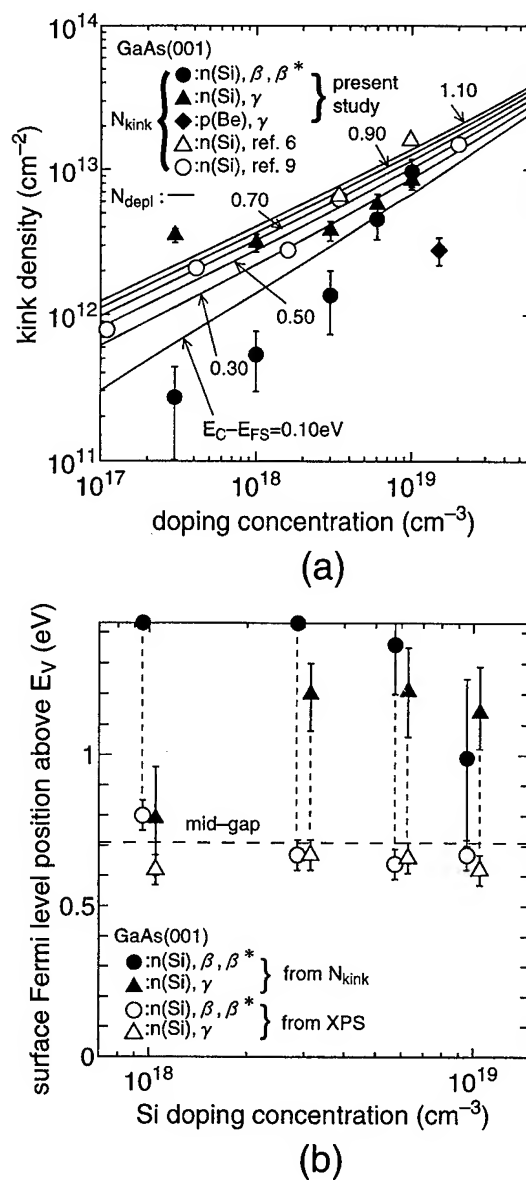


FIG. 11. (a) The calculated values of N_{depl} for various pinning positions and the experimentally observed kink density as a function of doping concentration of GaAs (2×4) surfaces. The kink densities experimentally obtained by the present study, Pashley and Haberern (Ref. 6) and Yamaguchi and Horikoshi (Ref. 9) are plotted by dots. (b) surface Fermi level position above E_V expected from the measured kink density after the kink-deep-acceptor model and by XPS in the present study. Doping values are shifted horizontally within doping measurement accuracy.

deep acceptors in the LT-GaAs case may be slightly smaller in the surface depletion region due to surface diffusion process, etc. We cannot entirely rule out such a possibility, although we feel it unlikely because post-growth annealing was done only for a short time of 3 min at the growth temperature in the Be-doped case. Further work is necessary to clarify this point.

Thus, we conclude that although the kink density has a strong correlation with Si doping on GaAs(001)-(2×4) surfaces, the observed kink density, which is strongly dependent on preparation conditions, is not always large enough to ex-

plain quantitatively the observed strong Fermi level pinning by the kink-deep-acceptor model.

The present result may seem to contradict the recent STS result by Pashley *et al.*,⁸ where a negative charge of $(1.2 \pm 0.4)e$ was directly detected near each kink. However, one should also remember that no discrete peaks corresponding to the existence of discrete deep acceptors were detected in the STS spectra. Instead, STS taken on the kink site showed either no gap states or broadly U-shaped continuous gap-state spectra similar to that shown in Fig. 10(b). This result is very difficult to understand by the kink-deep-acceptor model. Additionally, if we assume formation of acceptor type gap-state continuum by each kink through some unknown mechanism, the reported charge of $(1.2 \pm 0.4)e$ seems then to correspond to integration of continuous states up to E_C or above E_C and not to the mid-gap pinning position, judging from the empty state bias used. However, what matters for pinning is the integrated charge up to the pinning position, as is clear from Eq. (3), and this charge seems to be much smaller than $(1.2 \pm 0.4)e$.

Our tentative interpretation to explain the present result is that the major mechanism for strong Fermi level pinning in β -, β^* - and γ -phase surfaces observed in this study is due to disorder caused by missing As atoms on the As-dimer rows, holes, and steps, etc. that destroy two-dimensional order on the surface. They clearly exist as large densities in all the images which we took in our study, including β -phase samples. Although the ordered missing-dimer structure itself should be free of surface states according to Chadi,²² random disorder in the dimer array may introduce gap state continuum such as shown in Fig. 10(b). Namely, we assume that such surface disorder causes random scattering of electron wave and produces tail states. Of course, kinks are also the cause of such disorder, and may contribute to gap state continuum. Further theoretical and experimental work is obviously necessary to justify such an interpretation based on surface disorder. What we can conclude now is that kinks as simple discrete deep acceptors cannot explain the present result quantitatively.

B. InP (2×4) surface and possible reason for difference between GaAs and InP

XPS study showed that the Fermi level on Si-doped InP (2×4) surfaces is pinned at around $E_C - E_F = 0.3 - 0.6$ eV. Figure 12 shows the prediction of the kink-deep-acceptor model together with the measured kink densities.

Obviously, there is no correlation between kink density and Si doping concentration experimentally, and the observed number of kinks does not explain the observed Fermi level pinning. Thus, the kink-deep-acceptor model is not applicable to Fermi level pinning on the InP surface, either.

The observed difference between GaAs and InP surfaces with respect to correlation between Si doping and kink density is remarkable, and obviously requires explanation. In this connection, Yamaguchi and Horikoshi⁹ investigated the kinks on the (2×4) reconstructed InAs, and found no kink-related mid-gap Fermi level pinning. The result was ex-

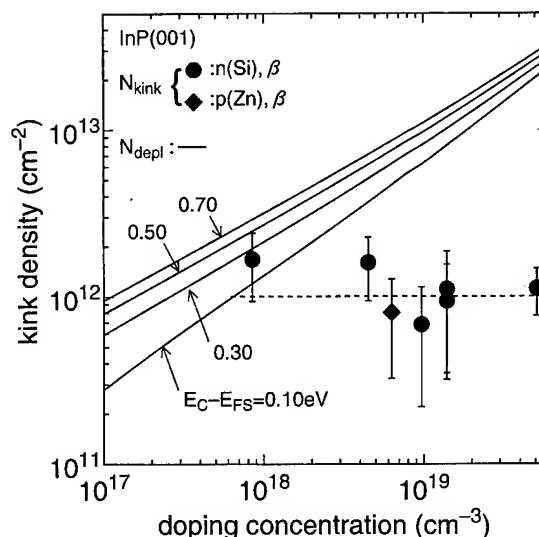


FIG. 12. The calculated values of N_{depl} for various pinning positions and the experimentally observed kink density in the present study as a function of doping concentration of InP (2×4) surfaces.

plained in terms of competition between energy loss by kink formation and energy gain by electron trapping from conduction band into the kink acceptor. This competition becomes more and more important in narrow band gap materials.

In the case of InP, no complication due to quantum effects, as discussed for InAs in Ref. 9, exists and their theory is basically the same with the original kink-deep-acceptor model of $N_{\text{kink}} = N_{\text{depl}}$ except that energy correction for the kink formation, E_{kink} , is included in Eq. (2). Thus, the theoretical curves remain the same in Fig. 12, and the values of the parameter $E_C - E_{\text{FS}}$ is reduced by the magnitude of E_{kink} , which is typically 0.2 eV according to Ref. 9.

Thus, these modified curves do not explain the observed constant kink density. In the case that the value of E_{kink} is unrealistically large, then no kink will be formed and this does not explain the observed rather high constant kink density either. Thus, it is difficult at present to provide a definitive explanation for the remarkable difference between GaAs and InP within the scope of this study. One possible explanation that we can suggest is that it is related to the difference in the actual microscopic structure of the (2×4) missing-dimer array between GaAs and InP, namely, the missing-dimer trench on GaAs (2×4) surface has 1 monolayer (ML) depth according to the latest observation.¹⁷ We also recently have confirmed this, and additionally we also have found that the missing-dimer trench on InP (2×4) surface has 0.5 ML depth.¹¹ Thus, when the Si atom is going to replace a group III atom on the (2×4) reconstructed surface, it should replace the one on the trench wall in the case of GaAs, whereas it should replace the one on the trench bottom on the InP surface. If the former is energetically unfavorable to the neighboring atoms as compared with the latter, it may lead to kink formation on GaAs, but not on InP. This is not to say that there is one Si atom at each kink site on GaAs. Instead, one kink once formed at one Si site may be

copied during crystal growth forming a kink without the presence of Si and such kinks may remain on the surface after the original Si atom itself is buried the bulk. Dynamic processes such as kink motion and recombination may also coexist during crystal growth and may complicate the kink dynamics. In any case, further work is necessary to clarify this point.

V. CONCLUSION

The relationship between kink defects and the Fermi level pinning on GaAs and InP(001)-(2×4) surfaces was studied in detail by STM and XPS. The main conclusions are listed below.

- (1) In Si-doped GaAs, the kink density increased with doping as previously found. However, actual density depended very much on the reconstruction phases and experimental conditions.
- (2) At high Si doping levels, Fermi level was strongly pinned below mid-gap, but the measured kink density was found to be not large enough to explain such pinning by the kink-deep-acceptor model, which assumes that each kink acts as a single discrete deep acceptor site. It is suggested that loss of two-dimensional order on the surface by various reasons including kinks may create surface states of continuous nature, and cause pinning.
- (3) In Si-doped InP, the kink density remained constant with the increase of Si doping, although the Fermi level was pinned above mid-gap. The result cannot be explained by the kink-deep-acceptor model either.

ACKNOWLEDGMENTS

The authors would like to thank N. Tsurumi for his assistance in the experiments. One of the authors (Y. I.) is financially supported by a Research Fellowship of the Japan Society for the Promotion of Science.

- ¹W. E. Spicer, Z. Liliental-Weber, E. Weber, N. Newman, T. Kendelewicz, R. Cao, C. McCants, P. Mahowald, K. Miyano, and I. Lindau, *J. Vac. Sci. Technol. B* **6**, 1245 (1988).
- ²V. Heine, *Phys. Rev.* **138**, 1689 (1965).
- ³J. Tersoff, *Phys. Rev. Lett.* **52**, 465 (1984).
- ⁴H. Hasegawa and H. Ohno, *J. Vac. Sci. Technol. B* **4**, 1130 (1986).
- ⁵J. M. Woodall and J. L. Freeouf, *J. Vac. Sci. Technol.* **19**, 794 (1981).
- ⁶M. D. Pashley and K. W. Haberern, *Phys. Rev. Lett.* **67**, 2697 (1991).
- ⁷M. D. Pashley, K. W. Haberern, R. M. Feenstra, and P. D. Kirchner, *Phys. Rev. B* **48**, 4612 (1993).
- ⁸M. D. Pashley, K. W. Haberern, and R. M. Feenstra, *J. Vac. Sci. Technol. B* **10**, 1874 (1992).
- ⁹H. Yamaguchi and Y. Horikoshi, *Phys. Rev. B* **53**, 4565 (1996).
- ¹⁰B. X. Yang, Y. Ishikawa, T. Ozeki, and H. Hasegawa, *Jpn. J. Appl. Phys.* **1** **35**, 1267 (1996).
- ¹¹Y. Ishikawa, T. Fukui, and H. Hasegawa, *Jpn. J. Appl. Phys.* **1** **36**, 1749 (1997).
- ¹²P. D. Kirchner, T. N. Jackson, G. D. Pettit, and J. M. Woodall, *Appl. Phys. Lett.* **47**, 26 (1985).
- ¹³E. A. Beam III, T. S. Henderson, A. C. Seabaugh, and J. Y. Yang, *J. Cryst. Growth* **116**, 436 (1992).
- ¹⁴E. Tokumitsu, *Jpn. J. Appl. Phys.* **1** **29**, L698 (1990).
- ¹⁵H. H. Farrell and C. J. Palmström, *J. Vac. Sci. Technol. B* **8**, 903 (1990).
- ¹⁶J. Zhou, Q. Xue, H. Chaya, T. Hashizume, and T. Sakurai, *Appl. Phys. Lett.* **64**, 583 (1994).
- ¹⁷A. R. Avery, C. M. Goringe, D. M. Holmes, J. L. Sudijono, and T. S. Jones, *Phys. Rev. Lett.* **76**, 3344 (1996).
- ¹⁸M. H. Hecht, *J. Vac. Sci. Technol. B* **8**, 1018 (1990).
- ¹⁹J. R. Waldrop, E. A. Kraut, C. W. Farley, and R. W. Grant, *J. Appl. Phys.* **69**, 372 (1991).
- ²⁰H. Ohno, H. Ishii, K. Matsuzaki, and H. Hasegawa, *J. Cryst. Growth* **95**, 397 (1989).
- ²¹W. Mönch, *Appl. Surf. Sci.* **92**, 367 (1996).
- ²²D. J. Chadi, *J. Phys. Soc. Jpn.* **47**, 1035 (1980).

Comparison of Si and GaAs/interfaces resulting from thermal and plasma oxidation

P. R. Lefebvre and E. A. Irene^{a)}

Department of Chemistry, University of North Carolina at Chapel Hill, Chapel Hill,
North Carolina 27599-3290

(Received 12 January 1997; accepted 10 April 1997)

X-ray photoelectron spectroscopy (XPS) analyses of oxides produced by thermal and plasma oxidation of GaAs show that electron cyclotron resonance (ECR) plasma oxidation favors what is thought to be the more desirable (from the point of view of electronic passivation) As⁺⁵ oxidation state while thermal oxidation favors the lower oxidation state, As⁺³. Thermal oxidation produces a Ga-rich oxide, whereas the ECR plasma oxides are nearly stoichiometric. Also, thermal oxidation removes As⁽⁰⁾ during the initial stage. XPS shows that thermal and ECR plasma Si oxides have different structures. *In situ* and real time ellipsometry studies indicate that thermal and ECR plasma oxidations yield different film growth kinetics for both GaAs and Si. In particular there are no strong substrate orientation effects for the ECR plasma oxidation of GaAs and Si in the initial stage of oxidation. The similarities in our Si and GaAs oxidation results enable models to emerge based on the dominance of the highly reactive plasma generated oxidant species for plasma oxidation and molecular oxygen for the thermal case. © 1997 American Vacuum Society. [S0734-211X(97)06504-9]

I. INTRODUCTION

The technology of the metal insulator semiconductor field effect transistor (MISFET) hinges on the electronic passivation of the semiconductor/insulator interface. An electronically passivated interface has a significantly smaller number of interface states in the band gap than the number of carriers needed for device operation. These states can arise from inadequate termination of the crystal lattice at the surface and are called dangling bonds or intrinsic interface states, as well as from defects and impurities, and are then referred to as extrinsic interface states. Electronic passivation of the Si surface is achieved through thermal oxidation, and the resulting interface state density is typically less than $10^{10} \text{ cm}^{-2} \text{ eV}^{-1}$, which is about 10^2 lower than the carrier concentration required for Si-based MISFET operation. It is known that electron cyclotron resonance (ECR) plasma oxidation produces good electrical quality oxides on Si.¹ Plasma oxidation of GaAs has also been previously reported.²⁻⁴ III-V semiconductors such as InP and GaAs offer materials property advantages over Si, such as increased carrier mobility and direct band gap, and may be used for microwave and optoelectronic applications. However, electronic passivation of GaAs is not fully resolved, and models have been proposed to explain the origin of the high density of interface states at III-V semiconductor/insulator interfaces.⁵⁻⁷ "Fixes" have been found using S,⁸ or running deionized water (RDIW) and UV light⁹ for instance, to reduce surface state levels. However, these solutions have not yet achieved wide acceptance due to processing and stability issues. A recent promising study¹⁰ involving *in situ* molecular beam

epitaxial deposition of various oxides shows an interface midgap density in the $10^{10} \text{ eV}^{-1} \text{ cm}^{-2}$ range, on *n* and *p*-type GaAs/Ga₂O₃ structures.

Previous work in our laboratory on InP (Ref. 11) led to a passivation method involving first mild ECR plasma oxidation to about 2 nm oxide thickness, to reduce intrinsic states, followed by the ECR plasma chemical-vapor deposition of SiO₂ to preserve the stoichiometry of the interface. This process did not produce excess P at the InP/oxide interface and resulted in good electrical characteristics of MIS structures. The present research follows from the InP studies. ECR plasma offers many advantages such as low ion energies, high plasma density, and the possibility to operate at room temperature, which makes it an important method for compound semiconductor processing.

In the present study, we report a comparison of the thermal and ECR plasma oxidation of GaAs and Si. We address such issues as the structural differences between the oxides obtained thermally, and with ECR plasma, the effect of the plasma conditions on the resulting oxides, and orientation effects on the kinetics of oxidation. First several GaAs cleaning procedures from the literature were compared using spectroscopic ellipsometry (SE), x-ray photoelectron spectroscopy (XPS), and atomic force microscopy (AFM), in order to find a process that leads to a reproducible and uniform starting surface. Also, the amount of As⁽⁰⁾ left after the cleaning, as well as produced during various oxidation processes has been monitored, since As⁽⁰⁾ is held partly responsible for poor electrical properties of GaAs MISFET devices. Thermal oxidations, which were included in this study mainly for the purposes of comparison with ECR plasma results, were performed on different GaAs orientations, at 420 °C, and on Si(100) at 500 °C and 800 °C. ECR plasma oxides were grown on various GaAs and Si substrates. Substrate bias and

^{a)}Electronic mail: GENE_IRENE@UNC.EDU

temperature effects were observed. The oxidations were monitored *in situ* and in real time with single wavelength ellipsometry (SWE), *in situ* with SE, and *ex situ* with XPS. Optical models of the oxides were established from SE data and checked with XPS.

II. EXPERIMENTAL PROCEDURES

The GaAs wafers used were polished single crystal semi-insulating undoped GaAs wafers with (100), (110), and (111) orientations. The (111) wafer type was determined to be a Ga terminated surface, using White and Roth's etchant.¹² The Si wafers were *p*-type *c*-Si, with (100), (110), and (111) orientation, and 1–2 ohm cm resistivity. Si surfaces were cleaned using a modified RCA procedure,¹³ with a final HF dip and deionized water rinse. As for GaAs cleaning, three different methods from the literature were tested, but all started with a degreasing step, involving sequential rinsing in electronics grade tetrachloroethylene, acetone and methanol, followed by RDIW for 2 min. Each step lasted 10 min, and was performed in an ultrasonic bath. The different chemical treatments applied after the degreasing step are listed below.¹⁴

Process 1: $\text{H}_3\text{PO}_4\text{:H}_2\text{O}_2\text{:H}_2\text{O}(1\text{:}1\text{:}10)$, room temperature, 4 s + RDIW rinse + $\text{HCl}\text{:MeOH}(1\text{:}1)$, room temperature, 2 min + final RDIW rinse.

Process 2: $\text{HCl}\text{:MeOH}(1\text{:}1)$, room temperature, 2 min + final RDIW rinse.

Process 3: $\text{HCl}\text{:H}_2\text{O}(1\text{:}1)$, room temperature, 2 min + final RDIW rinse.

After the last water rinse, the samples were blown dry with high purity N_2 , and then were transferred in air to the ECR plasma processing/SE chamber or to the XPS system. It

should be noted that other successful cleaning processes are reported in the literature, see for example Ref. 14, but we made no attempt to be exhaustive. Rather, we chose three of the simplest procedures to compare.

The GaAs thermal oxidations were performed at 420 °C in dry oxygen, in a cylindrical fused silica tube placed in a resistively heated furnace. This temperature was selected because we found that it was high enough to perform kinetics studies and low enough to preclude large amounts of production of elemental As. The flow of oxygen was 50 sccm. The Si thermal oxidations were performed at 500 and 800 °C.

The ECR plasma processing chamber/SE system is shown in Fig. 1. The plasma operates at 300 W and 2.45 GHz. The oxygen pressure was maintained at about 8×10^{-4} Torr during oxidation, and the sample was placed about 20 cm downstream from the plasma. The sample could be heated by a halogen lamp behind the sample stage, and a dc bias voltage between the sample and the plasma could be applied. The base pressure of chamber was around 5×10^{-8} Torr. The spectroscopic ellipsometer was a custom built rotating analyzer system with a spectral range of about 1.5 eV and 5.5 eV, and has been previously described.¹⁵ The SWE studies were carried out at 380 nm for GaAs, and 340 nm for Si, because at these energies, ellipsometry is very sensitive to an overlayer, but is relatively insensitive to temperature changes for GaAs¹⁶ or Si.¹⁷ Also, before SE measurements, the samples were cooled to room temperature. The thicknesses reported in this work that were derived from SWE measurements, were obtained using a trajectory method.¹⁷ The measurables in ellipsometry are the changes in amplitude, Ψ , and phase, Δ , of the *p* and *s* components of polarized light upon reflection from a sample surface. The complex reflection coefficient, $\rho = \tan \Psi e^{i\Delta}$, can be directly related to the dielectric function ϵ of a pure substrate through the relation

$$\epsilon = \epsilon_1 - i\epsilon_2 = \sin^2 \phi \left[\tan^2 \phi \left(\frac{1 - \rho}{1 + \rho} \right)^2 + 1 \right], \quad (1)$$

where ϕ is the angle of incidence, 71° in this case. For a nonperfect or film covered surface, Ψ and Δ yield a pseudodielectric function, $\langle \epsilon \rangle$, which is a convolution of the substrate and film(s) dielectric functions. The measured complex reflection coefficients, ρ_{exp} , were related to physical properties of the sample through model calculations using the Bruggeman effective medium approximation.¹⁸ To compare the measured $\rho(\rho_{\text{exp}})$ to that of an assumed model, ρ_{calc} , an unbiased estimator δ , is calculated from the relation¹⁹

$$\delta = \left[\frac{1}{N - P - 1} \sum (\rho_{\text{exp}} - \rho_{\text{calc}})^2 \right]^{1/2}, \quad (2)$$

where N is the number of wavelengths, P the number of unknown parameters, and ρ_{calc} is calculated from a literature database.²⁰ A minimizing procedure gives the best fit parameters which are film thickness and composition, at the 90% confidence level. To interpret the SWE measurements in terms of film thickness, we assumed a simple one layer

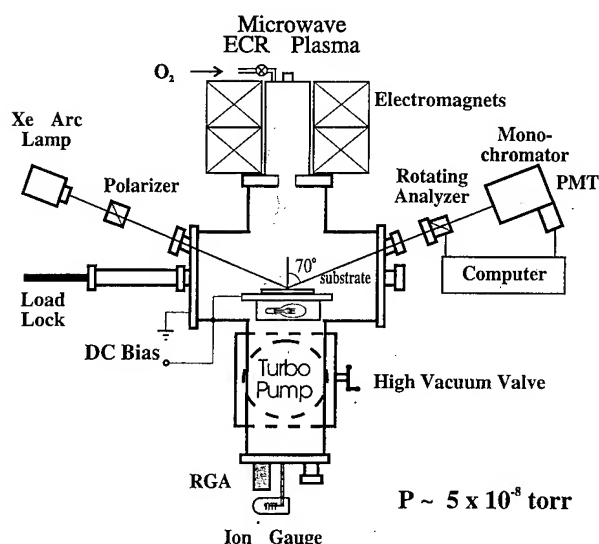


FIG. 1. Diagram of the ECR plasma processing chamber/SE integrated system.

model. Further discussion of modeling will be given below. Film thicknesses reported in this study were determined using ellipsometry.

The XPS spectra were taken using monochromatized Al $K\alpha$ radiation (1486.6 eV). Charging effects were eliminated using an electron flood gun, and the C 1s peak was kept around 285 eV. The spectra were taken at 80° and 20° take-off angles, which gives the possibility for depth profiling. The XPS raw data shown below have not been smoothed or convoluted. The ratios shown in Tables I, II, and IV were derived as follows:

$$\frac{Ga_{ox}}{As_{ox}} = \frac{\%Ga_{tot}}{\%As_{tot}} \times \frac{\text{area of } Ga_{ox} \text{ peak}}{\text{area of } As_{ox} \text{ peak}}, \quad (3)$$

where $\%Ga_{tot}$ and $\%As_{tot}$ are the total percentage of Ga and As detected in the sample. Similarly, the As ratios are given as

$$\frac{As^{+x}}{As^{+y}} = \frac{\text{area of } As^{+x} \text{ peak}}{\text{area of } As^{+y} \text{ peak}}. \quad (4)$$

The values in the tables are deduced from fitting, and a 20% error in the areas corresponds to a maximum of 50% error in the ratios. We believe such errors can occur when examining the $As^{(0)}$ peak. For GaAs, we analyzed mainly the As and Ga 3d core level spectra, but also the 2p spectra, which is another way of performing depth analysis, since the escape depths for the 3d and 2p electrons are significantly different²¹ (about 2.4 nm versus about 0.7 nm). For Si, the Si 2p peaks were used. In both cases, the O 1s window was also analyzed. The XPS data were fit using mixed Gaussian-Lorentzian functions, in order to perform semiquantitative analysis. The fitting includes the 3/2 and 5/2 contributions for the 3d peaks, with a ratio of 0.67, and a separation of 0.47 eV for Ga, and 0.69 for As. The 1/2 and 3/2 components of the Si 2p peaks are kept with a ratio of 0.5, and a separation of 0.59 eV. XPS data were not used to determine oxide thicknesses in this work. In this study, chemical etching of GaAs thermal oxide was also used with XPS, in order to better understand the depth composition of our oxides. A more detailed study²² of the depth composition of native oxide using a chemical depth profiling method in conjunction with XPS also showed a gradient of composition in the oxide, consistent with our results.

AFM was utilized to image surfaces from which rms roughness values were obtained.

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Cleaning

For Si, a modified RCA method¹³ was used, followed with a HF dip and deionized water rinse. This method is known to give reproducible surfaces, and SE shows that it leaves a $0.4 \text{ nm} \pm 0.1 \text{ nm}$ residual oxide layer. The SE data was analyzed using a single layer two component model. Our XPS measurements at 80° indicate that the residual oxide does not contain SiO_2 , but rather suboxides, hereafter referred to as SiO_x . The binding energy (BE) shift for

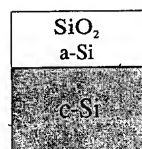
TABLE I. XPS data at 80° for cleaning processes 1 through 3, for a GaAs (100) sample. For the details of the cleaning processes, see the text. The table gives the Ga to As oxide ratio, the As^{+5} to As^{+3} oxide ratio, and the ratio of the $As^{(0)}$ to that of As in GaAs, derived from the XPS 3d peaks. The energies are in eV.

Process	Ga_{ox}/As_{ox} 3d	As^{+5}/As^{+3} 3d	$As^{(0)}/As_{subst}$ 3d
1	0.3	0	0.25
2	0.4	0	0.10
3	0.25	0	0.18

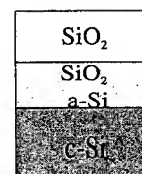
SiO_x , measured from substrate Si, is 1.05 eV, which corresponds to a +1 oxidation state for Si.²³ For GaAs, three different cleaning processes described above were compared using SE and XPS. All of the processes left a residual overlayer presumably formed during air exposure after the last water rinse, which is anticipated since it has been reported²⁴ that the HCl treated GaAs surface is very reactive in air. Our intention in the "cleaning" phase of this work was not to achieve a bare GaAs surface, but rather a reproducible surface with a low concentration of electronic defects for potential MIS applications. Therefore, close attention was paid to the $As^{(0)}$ in the residual oxide layer. The thickness of the overlayer, after the various method tested, was $1.2 \text{ nm} \pm 0.2 \text{ nm}$, determined from SE data, and analyzed using a single film two component model. Table I summarizes the XPS fitting results obtained from methods 1 through 3. For the Ga 3d peak, two components separated by $1.05 \pm 0.05 \text{ eV}$ were observed. They correspond to the substrate Ga appearing at a lower BE, and to Ga^{+3} for the other. As for the As window, the oxide peak has only one component, As^{+3} , separated from the substrate by $3.37 \pm 0.03 \text{ eV}$. An additional compo-

Si models

MODEL 1

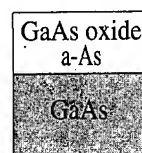


MODEL 2



GaAs models

MODEL 1



MODEL 2

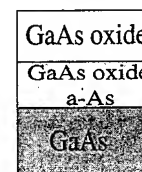


FIG. 2. One and two film models used to fit the SE data, for Si and GaAs.

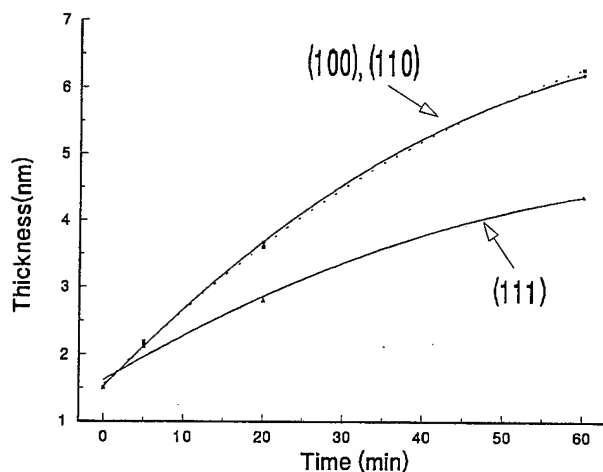


FIG. 3. Substrate orientation effect on the thermal oxidation of GaAs at 420 °C. The line is a polynomial fit to the data, and is not related to a specific model of oxidation.

nent is observed that is shifted by about 0.6 eV relative to the As substrate peak and is attributed to elemental arsenic, $\text{As}^{(0)}$, which is known to accumulate at the GaAs/oxide interface.²⁵ Table I shows that the residual oxide is not stoichiometric, it is As rich, it does not contain As^{+5} , and that different cleaning methods produce different interfaces. The process using H_3PO_4 seems to leave the most $\text{As}^{(0)}$ at the interface, and the $\text{HCl}:\text{MeOH}$ (1:1) mixture seems to leave less $\text{As}^{(0)}$ than does $\text{HCl}:\text{H}_2\text{O}$ (1:1). Another study²⁶ shows

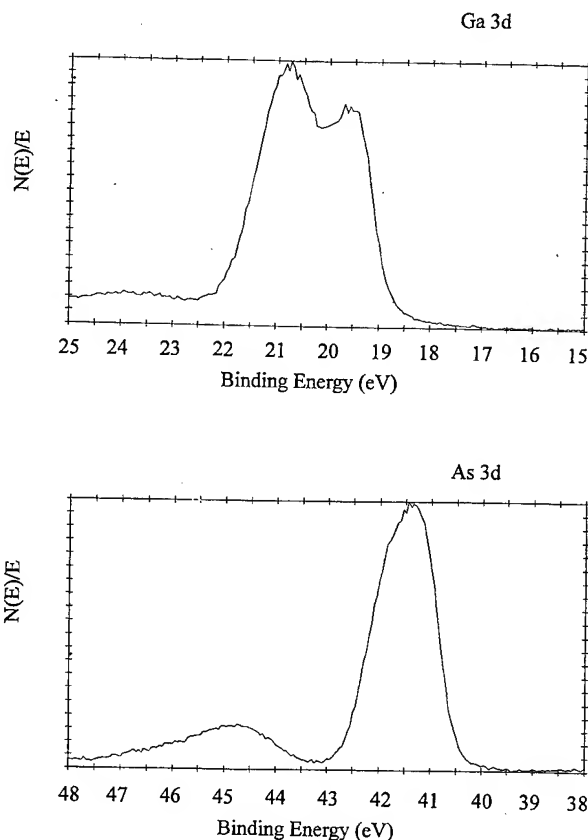
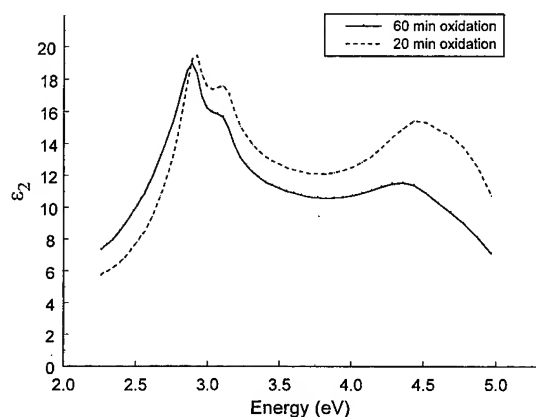


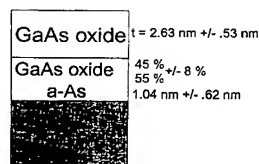
FIG. 5. Representative Ga and As 3d XPS peaks for a GaAs sample thermally oxidized at 420 °C for 20 min.



(a)

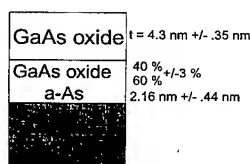
20 min

60 min



$\delta = 0.0048$

(b)



$\delta = 0.0046$

FIG. 4. (a) Imaginary part of the pseudodielectric function, $\langle \epsilon_2 \rangle$, for GaAs samples thermally oxidized at 420 °C for 20 min and 60 min, respectively, and (b) SE modeling data for the same samples, deduced from (a).

that a $\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2:\text{H}_2\text{O}$ (5:1:1) mixture leaves more $\text{As}^{(0)}$ at the interface than a concentrated HCl (50%) dip. Procedure 2 was therefore chosen as our cleaning method for this study, since it leads to reproducible surfaces, and AFM indicates smooth surfaces with a rms value of about 0.4 nm.

B. Thermal oxidation

Si thermal oxidation kinetics measurements were not re-done, since the results from extensive studies are available.²⁷ It is known that in the initial oxidation regime (<10 nm) the (110) Si surface oxidizes faster than the (111), which in turn oxidizes faster than (100); this order parallels the areal density of Si atoms.

The thermal oxidation of the (100), (110), and (111) GaAs orientations was followed with SE. Figure 2 displays the optical models that were used in this study, in order of complexity. The simplest models involve a single film with one component, and this works well with thick SiO_2 layers, where the contribution of the interface is negligible. However for our thickness range, we need to use a single film two component model, or a two film model if the interface is large enough, or to model interface roughness. The optical absorption of the films increases with the content of elemental As, or $a\text{-Si}$, as well as with the roughness of the interface. By the same token, $\text{As}^{(0)}$, or $a\text{-Si}$, can be used in the modeling as constituents that increase n and/or k of a film, for

TABLE II. XPS data at 80° for a GaAs (100) sample oxidized thermally for 20 min. (a) shows the BE differences between the Ga or As substrate peak and the corresponding oxides for the 3d windows, the FWHM, shown in square brackets, of the oxide peaks, and the BE differences between the O 1s peak and the various 3d oxide peaks. (b) shows the stoichiometry of the oxides, as well as the ratio of As⁽⁰⁾ to As of the substrate. The energies are in eV.

	Ga ⁺³ -Ga 3d	As ⁺³ -As 3d	As ⁺⁵ -As 3d	O _{1s} -Ga ⁺³	O _{1s} -As ⁺³	O _{1s} -As ⁺⁵	Ga _{ox} /As _{ox} 3d	Ga _{ox} /As _{ox} 2p	As ⁺⁵ /As ⁺³ 3d	As ⁽⁰⁾ /As _{subst} 3d
Before	1.33	3.55	4.90	510.9	486.95	485.6	7	2.5	0.9	0.015
H ₂ O dip	[1.15]	[1.45]	[1.61]							
After H ₂ O dip	1.31 [1.15]	3.46 [1.57]	4.85 [1.79]	510.95	487.1	485.7	14	5.8	1.25	0.024

example if the density of the measured film is higher than that used for the data base. For the GaAs oxide, the available data base is that of an anodic oxide, which contains equal amounts of As⁺³ and Ga⁺³ oxides. To decide which is the best model for a specific situation, we relied on the value of δ , as defined in Eq. (2), as well as on physical facts deduced from our XPS data and/or from the literature.

Figure 3 shows thermal oxidation results for three GaAs orientations. A single layer model with two components was used in order to determine the thicknesses. It is seen that the (100) and (110) surfaces exhibit essentially the same behavior while the (111) A samples oxidize more slowly. A (111)A surface has Ga surface atoms attached to three As atoms below the surface plane, and the valency of Ga is satisfied. This situation is not favorable to oxidation, because electrons are not readily available. On the other hand, the (110) orientation, with both As and Ga atoms at the surface, oxidizes faster than the (111)A orientation. The (100) orientation also oxidizes faster and this is possibly attributable to the fact that on this surface each atom is attached to two atoms in the plane below yielding uniform but weaker bonding than the (111) surface. These results are consistent with published etch rate studies¹⁴ on various GaAs orientations. Thus for GaAs thermal oxidation, the rates of oxide formation appear to be dominated by the chemical environment at the GaAs surface rather than simply by the atomic density. Figure 4 displays representative SE data in terms of $\langle \epsilon_2 \rangle$ for GaAs samples thermally oxidized at 420 °C for 20 and 60 min, as well as the modeling results using model 2 for the same samples. Model 2 gave better results than model 1 in terms of δ . It is seen that during the thermal oxidation of GaAs, both the internal (interfacial) and external (outer) layers grow. In order to better understand the model, we performed XPS measurements on a sample oxidized for 20 min at 420 °C. Figure 5 shows representative Ga and As 3d spectra for this oxide. The sample was dipped in running deionized water for a minute in order to dissolve the top part of the oxide layer, and reanalyzed with XPS. The fitting results are shown in Table II. It has been argued²⁸ that oxidation of GaAs does not produce separate phases of Ga₂O₃, As₂O₃, or As₂O₅, but rather a single phase nonstoichiometric oxide with mixed bond types. For this reason, the individual oxides are not identified in our analysis. The XPS results confirm that thermal oxidation produces Ga rich oxides,²⁵ the ratio of Ga to As oxide being significantly greater than 1. Extensive XPS studies²⁹ on thermally grown GaAs oxides also show

that As⁺³ is the dominant form of As oxide, even though As⁺⁵ was detected. In our case, the presence of such a relatively large amount of As⁺⁵ might be due to the fact that our oxidation temperature was low, and our oxidation time short. Comparison of the ratios of Ga to As oxides using the 3d and 2p peaks indicates that the outer part of the oxide is richer in As than the substrate/oxide interface. This fact is confirmed by the H₂O dip, which mostly eliminates arsenic oxides. At our oxidation temperature, the arsenic oxides, As₂O₃ and As₂O₅ should evaporate, and yet we detect oxidized forms of As. This confirms previous studies²⁸ that suggest that the GaAs thermal oxide produced at this temperature is a single phase structure. As⁽⁰⁾ is observed at the interface, but its amount has been reduced compared to that of a freshly cleaned sample. The As⁽⁰⁾ pileup at the interface has been explained in two ways: first through the ternary condensed phase diagram³⁰ which predicts that GaAs and As₂O₃ react to give GaAs and As⁽⁰⁾; and second using Wilmsen's model for oxide growth³¹ which predicts that during the early stage of oxidation, the As oxides evaporate, leaving a layer of Ga rich oxide behind, through which further outdiffusion of As is difficult. It is likely that we do not observe much As⁽⁰⁾ at the interface, because our reaction temperature is low, and the reaction time is short. The apparent increase in the As⁽⁰⁾/As_{subst} ratio observed after the H₂O dip cannot be interpreted since it is within the error of the experimental method. On the other hand, our SE data shows that the As⁽⁰⁾ content of our oxides is high. However this may be a modeling artifact, considering that the oxide near the GaAs/oxide interface is Ga rich, and therefore does

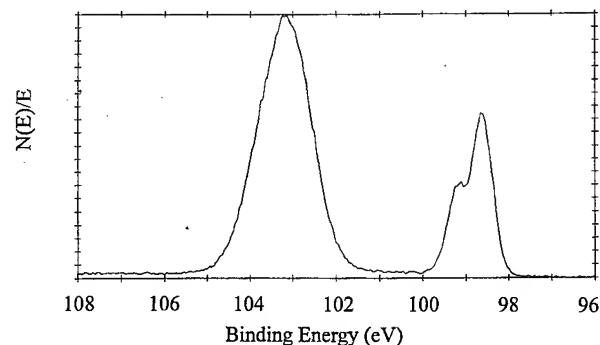
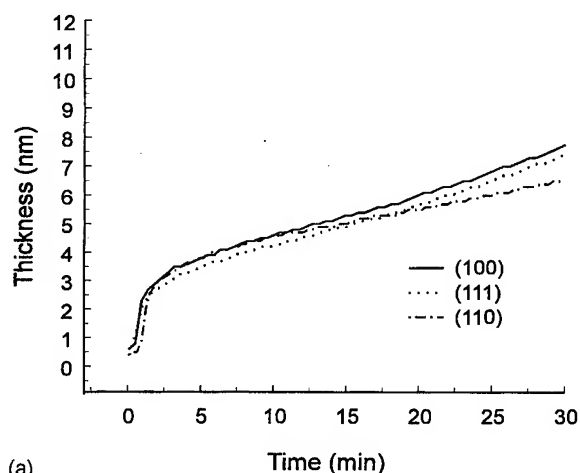


FIG. 6. Representative Si 2p XPS peak for a Si sample oxidized at 800 °C for 30 min.

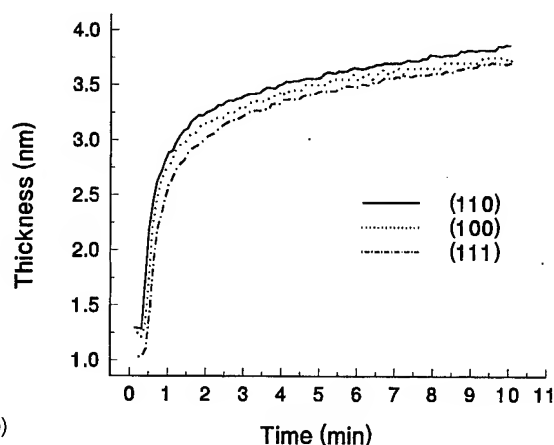
TABLE III. XPS data at 80° for clean Si, and for Si thermal oxides grown at 800 °C for 30 min and 3 h, and 500 °C for 14 h. The table displays the O 1s peak position as well as the BE differences between the oxides detected and the substrate peak. The FWHMs are shown in square brackets. The energies are in eV.

	O 1s	Si ⁺¹ -Si	Si ⁺² -Si	Si ⁺³ -Si	Si ⁺⁴ -Si	O 1s-Si ⁺⁴
Clean Si	532.34 [1.88]	1.04 [0.63]
800 °C 30 min	532.68 [1.34]	1.01 [0.63]	4.37 [1.21]	429.78
800 °C 3 h	532.9 [1.25]	4.31 [1.15]	429.75
500 °C 14 h	532.45 [1.46]	1.01 [0.70]	1.86 [0.74]	2.62 [1.08]	4.13 [1.26]	429.86

not correspond to our data base used for model fitting. Therefore the *a*-As in the model might be needed in order to account for a denser part of the oxide for example, and indeed Ga₂O₃ is more dense than any As containing oxide.³²



(a)



(b)

FIG. 7. Substrate orientation effect on the ECR plasma oxidation of (a) Si with a substrate bias of +30 V at floating temperature and (b) GaAs with a substrate bias of +10 V at floating temperature. The plasma power was 300 W.

Figure 6 shows the Si 2p spectrum for a Si sample oxidized at 800 °C for 30 min. Table III displays the XPS fitting results in the present study for a clean Si substrate, and for three Si thermal oxides. Two samples were grown at 800 °C for 30 min and 3 h, respectively, and the other at 500 °C for 14 h. The thicknesses of these oxides, determined by SE, are 3.9 nm and 15.6 nm for the 800 °C samples, respectively, and 2.4 nm for the sample oxidized at 500 °C. SiO₂ is not observed on our freshly cleaned sample at 80°, and only a component shifted by 1.04 eV is present. Our data at 25° (not shown), indicates a small component shifted by 3.82 eV. This is probably SiO₂, considering that the BE shift between the substrate and the SiO₂ decreases with decreasing oxide thickness,³³ and the effect of the oxide thickness on the BE shift ceases at about 5 nm. Our results show that the oxidation at 500 °C leaves suboxides even after 14 h of oxidation, and that the full width at half maximum (FWHM) of the SiO₂ peak is greater than that of 800 °C oxide. The O 1s peak of the oxide grown at 500 °C is also greater than that of the 800 °C ones. These observations indicate that the process temperature affects the structure of the oxides, and this will be discussed further when we compare the thermal oxides with the ECR plasma grown oxides.

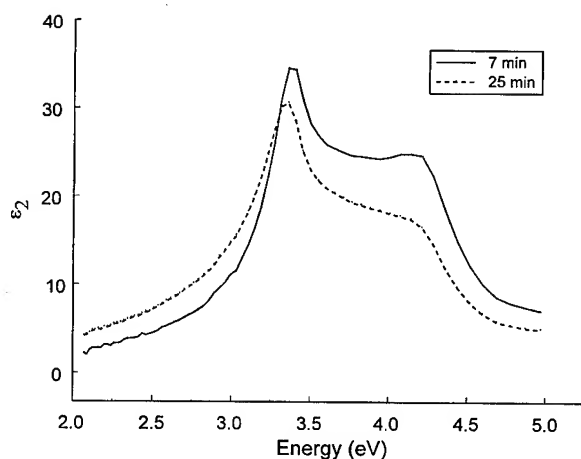


FIG. 8. Imaginary part of the pseudodielectric function (ϵ_2), for ECR plasma grown oxides on Si, for 7 min and 25 min with a substrate bias of +30 V.

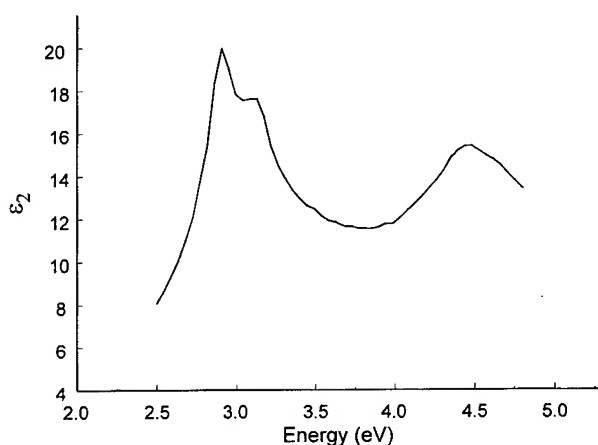
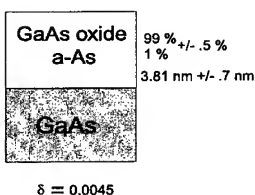


FIG. 9. Imaginary part of the pseudodielectric function, $\langle \epsilon_2 \rangle$, for an ECR plasma grown oxide on GaAs, 10 min with a substrate bias of +10 V.

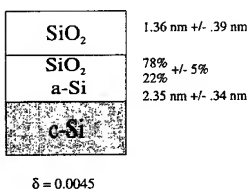
C. ECR plasma oxidation

Figure 7 shows the ECR plasma oxidation kinetics data for Si and GaAs, for three orientations as obtained from SWE measurements. SE data were also taken at different stages of the growth, to confirm the SWE results. It is seen that, for the thickness range studied, the Si and GaAs orientation has no significant effect on the kinetics of oxidation, which is contrary to the thermal results. In fact, if there is any effect on Si, it is interesting to note that it is the (110) surface that oxidizes the slowest and which for thermal oxidation is the fastest oxidizing orientation.²⁶ The effect of substrate bias and temperature on the kinetics of ECR plasma oxidation has already been studied for Si³⁴ and GaAs.³⁵ It

ECR GaAs 10 min



ECR Si 7 min



25 min

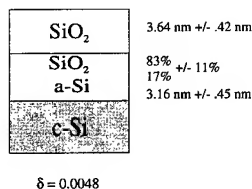


FIG. 10. SE modeling data for ECR plasma oxides grown on GaAs, 10 min with a substrate bias of +10 V, and on Si, for 7 min and 25 min, respectively, with a substrate bias of +30 V.

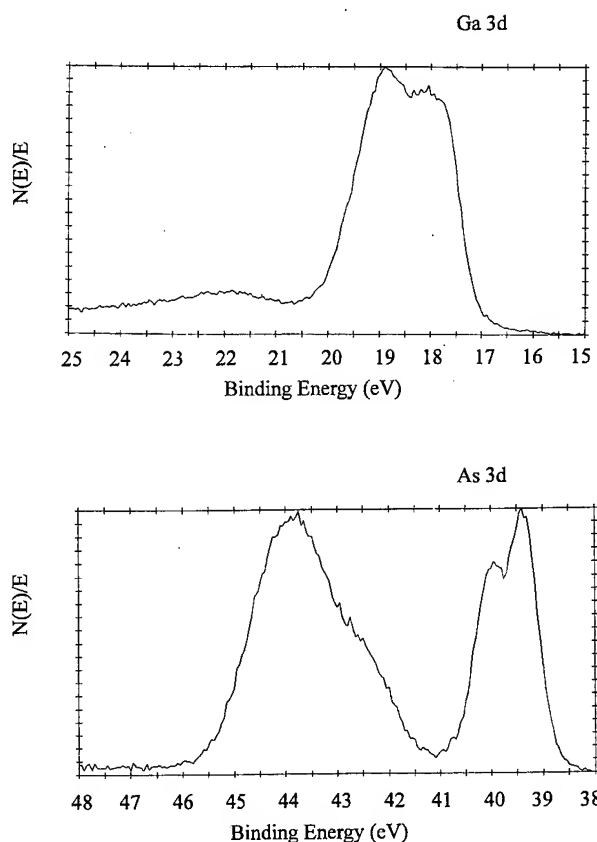


FIG. 11. Ga and As 3d XPS peaks for a GaAs sample oxidized at a plasma power of 300 W for 10 min, with a +10 V substrate bias.

was shown that in both cases a positive substrate bias, as well as higher temperature, enhances the oxidation rate, and also that the initial regime of oxidation, that is the first 5 min, is independent of substrate bias.

Figures 8 and 9 show the SE data in terms of $\langle \epsilon_2 \rangle$ for ECR plasma grown oxides on Si and GaAs, and Fig. 10 displays modeling results for these samples, deduced from the previously shown SE data. Model 2, which includes an interfacial layer, was found not to apply to ECR plasma oxides on GaAs where the interfacial layer thickness was found to be zero. Figure 10 shows representative modeling results for a GaAs sample oxidized for 10 min with a substrate bias of +10 V, and similar results were found for various biases and oxidation times. This modeling indicates that the GaAs ECR plasma oxides are more uniform than the thermal oxides. This result is likely due to temperature, since we have shown in our previous study³⁵ that annealing a GaAs ECR plasma grown oxide sample in vacuum significantly changed the structure of the oxide, to the point where model 2 had to be used in order to analyze the SE data. Model 2 was used for Si, and indicates that there is an interfacial layer between the substrate and the oxide. This layer increases slightly with oxidation time.

Figure 11 shows representative Ga and As 3d spectra for a GaAs sample, oxidized at a plasma power of 300 W for 10 min, with a +10 V substrate bias. Table IV summarizes the

TABLE IV. XPS data at 80°, unless stated otherwise, for a GaAs (100) sample oxidized by ECR plasma for 10 min. The plasma power was 300 W, the substrate bias was +10 V, and the substrate was oxidized at floating temperature. (a) The BE differences between the Ga or As substrate peak and the corresponding oxides for the 3d windows, the FWHM of the oxide peaks in square brackets, and the BE differences between the O 1s peak and the various 3d oxide peaks. (b) The stoichiometry of the oxides, as well as the ratio of As⁽⁰⁾ to As of the substrate. The energies are in eV.

(a)	Ga ⁺³ -Ga 3d	As ⁺³ -As 3d	As ⁺⁵ -As 3d	O _{1s} -Ga ⁺³	O _{1s} -As ⁺³	O _{1s} -As ⁺⁵
+10 V	1.10	3.16	4.62	511.30	487.82	486.2
10 min	[1.15]	[1.45]	[1.65]
(b)	Ga _{ox} /As _{ox} 3d	Ga _{ox} /As _{ox} 3d, 20°	As ⁺⁵ /As ⁺³ 3d	As ⁽⁰⁾ /As _{subst} 3d		
+10 V	1.1	1.1	3.50	0.1		
10 min						

XPS fitting results for that sample. The BE shifts between the substrate and oxide peaks are smaller for the ECR plasma grown oxides than for the thermal oxides. Also the shifts between the O 1s and the oxide peaks are larger for the ECR plasma grown oxides than for the thermal oxides, which indicates that the oxide structures are different. The stoichiometry of the oxides is also different: the plasma oxide is nearly stoichiometric, with a Ga to As oxide ratio close to unity, and contains more As⁺⁵ than the thermal oxide. These results agree with those of Lu *et al.*⁴ Figure 12 displays a representative Si 2p spectrum for a Si ECR plasma oxide grown for 7 min at a plasma power of 300 W, and with a +30 V substrate bias. Table V summarizes the XPS fitting results for Si ECR plasma oxides grown for 7 and 40 min, respectively, at a plasma power of 300 W, and with a +30 V substrate bias. The thicknesses of these oxides, determined by SE, are 3.7 nm for the 7 min sample, and 13.2 nm for the 40 min sample. The thin ECR plasma grown oxide shows a Si⁺² component that was not detected in the thermal sample. The FWHM of the O 1s and SiO₂ peaks are larger for the ECR plasma oxides, than for the thermal ones. It has been argued²³ that bond length/angle distribution affects the XPS

oxide peaks for Si, and that stress leads to broader oxide peaks. It is possible that our plasma oxide, grown at floating temperature has a more stressed interface than the thermally grown oxide. To confirm that our XPS results were temperature related, we grew a thermal oxide at 500 °C and analyzed it using XPS. The fitting results are reported in Table III. The FWHM are larger than in the 800 °C case, which corroborates our assertion of a temperature related phenomenon.

Yamasaki *et al.*³ investigated the transport of O₂ in GaAs oxide during dry plasma oxidation. They concluded that the mass transport in the oxide was due to ion drift. They also suggested that far from the plasma discharge, neutral oxygen comprises the dominant oxidant, and that a fraction attaches to the substrate surface, and is ionized through electron bombardment. Also, it has been argued³⁶ that O₂ decomposes to O and O⁻, which migrates through the oxide at positive bias. This corroborates our previous results³⁵ that show that the oxide growth rate is proportional to the positive bias value. Also in a previous study³⁷ we have shown that inserting a shutter between the plasma discharge and the sample slows the oxidation. The shutter stops direct impingement of electrons onto the substrate, and much of the UV irradiation from the plasma, which may also account for its effect on oxidation rate. Carl *et al.*³⁸ performed a detailed study on the ECR plasma oxidation of Si. Their kinetics data, although in

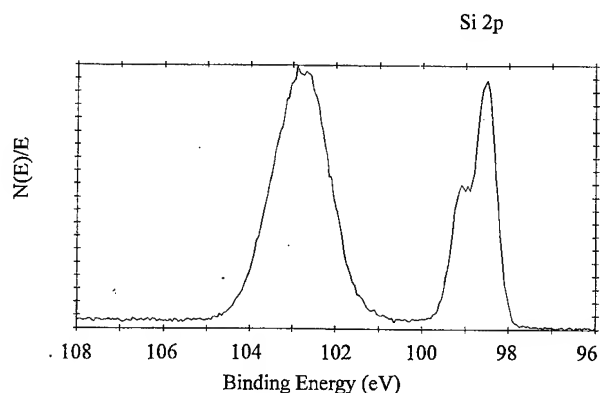


FIG. 12. Si 2p XPS peak for a Si ECR plasma grown oxide for 7 min at plasma power of 300 W, and substrate bias of +10 V.

TABLE V. XPS data at 80° for ECR plasma grown oxides on Si, for 7 and 40 min, respectively. The plasma power was 300 W, the substrate bias was +30 V, and the substrate was oxidized at floating temperature. The table displays the O 1s peak position as well as the BE differences between the oxides detected and the substrate peak. The FWHMs are shown in square brackets. The energies are in eV.

	O 1s [FWHM]	Si ⁺¹ -Si [FWHM]	Si ⁺² -Si [FWHM]	Si ⁺³ -Si [FWHM]	Si ⁺⁴ -Si [FWHM]	O 1s-Si ⁺⁴
ECR	532.15	1.03	2.09	...	4.08	429.70
7 min	[1.51]	[0.62]	[1.0]		[1.34]	
ECR	532.34	4.32	429.79
40 min	[1.50]				[1.37]	

a much thicker range, agree with ours. On the other hand, based on Fourier transform infrared spectroscopy, XPS, and Auger spectroscopy, they observed no chemical differences between thermal and plasma oxide. This might be due to the fact that their oxidations were done at higher temperatures than ours.

IV. CONCLUSIONS

A combination of SE and XPS showed that the oxide/substrate interfaces produced by thermal and ECR plasma oxidation of Si and GaAs are different. For GaAs, ECR plasma oxidation produces a more uniform and nearly stoichiometric oxide, with a higher As⁺⁵ content than the thermal oxide. For Si, it is not so much the suboxides content that distinguishes the oxides produced thermally and with ECR plasma, but the SiO₂ structure. The oxide produced by ECR plasma may be more stressed, and this appears to be related to the low temperature process. It has also been shown using SWE and SE data that the orientation effect seen in thermal oxidation for both Si and GaAs disappears in the initial stage of ECR plasma oxidation. Also, while for Si thermal oxidation, the order of oxidation rate follows the areal density of Si atoms on the surface, for GaAs, the rates of oxide formation appear to be dominated by the chemical environment at the GaAs surface. We have also shown that different cleaning methods lead to different starting surfaces in terms of As⁽⁰⁾ content. To more precisely define the GaAs interfaces, ultraviolet photoelectron spectroscopy measurements on the various oxides would be useful. Also, the various GaAs/oxide interfaces need to be electrically characterized, to determine the potential electronics significance of the different chemical interfaces.

ACKNOWLEDGMENTS

This research was partially supported by a grant from the National Science Foundation (NSF). The authors acknowledge Professor G. Hollinger of the Ecole Centrale de Lyon, France, for useful discussions on the interpretation of our XPS data.

- ¹Y. Z. Hu, J. Joseph, and E. A. Irene, Appl. Phys. Lett. **59**, 1353 (1991).
- ²R. P. H. Chang and A. K. Sinha, Appl. Phys. Lett. **29**, 56 (1976).
- ³K. Yamasaki and T. Sugano, J. Vac. Sci. Technol. **17**, 959 (1980).
- ⁴Z. Lu, M. T. Schmidt, R. M. Osgood, Jr., W. M. Holber, and D. V. Podlesnik, J. Vac. Sci. Technol. A **9**, 1040 (1991).
- ⁵H. Hasegawa, L. He, H. Ohno, T. Sawada, T. Haga, Y. Abe, and H. Takahashi, J. Vac. Sci. Technol. B **5**, 1097 (1987).
- ⁶J. M. Woodall and J. F. Freouf, J. Vac. Sci. Technol. **19**, 794 (1981).
- ⁷W. E. Spicer, Z. Liental-Weber, E. Weber, N. Newman, T. Kendelewicz,

- R. Cao, C. McCants, P. Mahohald, K. Miyano, and I. Lindau, J. Vac. Sci. Technol. B **6**, 1245 (1988).
- ⁸E. Yablonovitch, C. J. Sandroff, R. Bhat, and T. Gmitter, Appl. Phys. Lett. **51**, 439 (1987).
- ⁹S. D. Offsey, J. M. Woodall, A. C. Warren, P. D. Kirchner, T. I. Chappel, and G. D. Pettit, Appl. Phys. Lett. **48**, 475 (1986).
- ¹⁰M. Hong, M. Passlack, J. P. Mannaerts, J. Kwo, S. N. G. Chu, N. Moriya, S. Y. Hou, and V. J. Fratello, J. Vac. Sci. Technol. B **14**, 2297 (1996).
- ¹¹Y. Z. Hu, M. Li, E. A. Irene, M. Rowe, and H. C. Casey, Jr., Appl. Phys. Lett. **63**, 1113 (1993).
- ¹²J. G. White and W. C. Roth, J. Appl. Phys. **30**, 947 (1959).
- ¹³W. Kern and D. A. Puotinen, RCA Rev. **31**, 187 (1970).
- ¹⁴S. D. Mukherjee and D. W. Woodard, *Gallium Arsenide, Material, Devices and Circuits*, edited by M. J. Howes and D. V. Morgan (Wiley, New York, 1985), Chap. 4.
- ¹⁵J. W. Andrews, Y. Z. Hu, and E. A. Irene, Proc. SPIE **1188**, 162 (1989).
- ¹⁶D. E. Aspnes, W. E. Quinn, and S. Gregory, Appl. Phys. Lett. **56**, 2569 (1990).
- ¹⁷E. A. Irene, Thin Solid Films **233**, 96 (1993).
- ¹⁸D. E. Aspnes and A. A. Studna, Appl. Opt. **14**, 220 (1975).
- ¹⁹D. E. Aspnes, J. B. Theeten, and F. Hottier, Phys. Rev. B **20**, 3992 (1979).
- ²⁰D. E. Aspnes (private communication).
- ²¹H. Grant and W. Monch, Surf. Sci. **105**, 217 (1981).
- ²²F. J. Grunthaner, P. J. Grunthaner, R. P. Vasquez, B. F. Lewis, J. Maserjian, and A. Madhukar, J. Vac. Sci. Technol. **16**, 1443 (1979).
- ²³F. J. Grunthaner and P. J. Grunthaner, *Chemical and Electronic Structure of SiO₂/Si Interface*, Materials Science Reports (North Holland, Amsterdam, 1986), Vol. 1, pp. 65–160.
- ²⁴D. E. Aspnes, Thin Solid Films **89**, 249 (1982).
- ²⁵W. Wilmsen, *Physics and Chemistry III-V Compound Semiconductor Interfaces*, edited by C. W. Wilmsen (Plenum, New York, 1985).
- ²⁶J. Olivier, J. P. Landesman, and F. Wyczisk, presented at International Symposium on GaAs and Related Compounds, New Jersey, 1990 (unpublished).
- ²⁷E. A. Irene, CRC Crit. Rev. Solid State Mater. Sci. **14**, 175 (1988).
- ²⁸G. Hollinger, R. Skeyta-Kabbani, and M. Gendry, Phys. Rev. B **49**, 159 (1993).
- ²⁹G. P. Schwartz, G. J. Gualtieri, G. W. Kammlott, and B. Schwartz, J. Electrochem. Soc. **126**, 1737 (1979).
- ³⁰G. P. Schwartz, Thin Solid Films **103**, 3 (1983).
- ³¹C. W. Wilmsen, R. W. Kee, and K. M. Geib, J. Vac. Sci. Technol. **16**, 1434 (1979).
- ³²*CRC Handbook of Chemistry and Physics*, 71 ed., edited by D. R. Lide (Chemical Rubber, Boca Raton, FL, 1990-1991).
- ³³F. J. Grunthaner, P. J. Grunthaner, R. P. Vasquez, B. F. Lewis, J. Maserjian, and A. Madhukar, Phys. Rev. Lett. **43**, 1683 (1979).
- ³⁴J. Joseph, Y. Z. Hu, and E. A. Irene, J. Vac. Sci. Technol. B **10**, 611 (1992).
- ³⁵P. R. Lefebvre and E. A. Irene, *State of the Art Program on Compound Semiconductors XXIV*, edited by F. Ren, S. J. Pearton, and W. Platschen (The Electrochemical Society, Pennington, NJ, 1996).
- ³⁶J. Joseph, Y. Z. Hu, and E. A. Irene, *The Physics and Chemistry of SiO₂ and the Si-SiO₂ Interface 2*, edited by C. R. Helms and B. E. Deal (Plenum, New York, 1993).
- ³⁷P. R. Lefebvre and E. A. Irene (unpublished). The same was observed for InP: see Ref. 11.
- ³⁸D. A. Carl, D. W. Hess, M. A. Lieberman, T. D. Nguyen, and R. Gronsky, J. Appl. Phys. **70**, 3301 (1991).

Wet oxidation of AlAs films under ultrahigh vacuum conditions

W. J. Mitchell^{a)}

Center for Quantized Electronic Structures (QUEST), University of California, Santa Barbara, California 93106

C.-H. Chung

Department of Chemical Engineering, University of California, Santa Barbara, California 93106

S. I. Yi

Center for Quantized Electronic Structures (QUEST) and Department of Chemical Engineering, University of California, Santa Barbara, California 93106

E. L. Hu

Center for Quantized Electronic Structures (QUEST), University of California, Santa Barbara, California 93106

W. H. Weinberg

Center for Quantized Electronic Structures (QUEST) and Department of Chemical Engineering, University of California, Santa Barbara, California 93106

(Received 13 January 1997; accepted 3 April 1997)

The initial stages of oxidation of AlAs(001) (using D₂O as the oxidant) have been investigated using Auger electron spectroscopy and temperature-programmed desorption. We have found that molecularly adsorbed water on AlAs(001) has two competing reaction pathways available: either desorption back into the gas phase, or dissociation resulting in aluminum oxide, aluminum hydroxide, and arsenic hydride. Recombination of the arsenic hydride produces arsine, which desorbs and depletes arsenic within the oxide film, a process which is shown to enhance the oxide growth. By identifying the various reaction steps that occur (with annealing) after the low-temperature adsorption of water on AlAs(001), we are able to propose a mechanism for the initial stages of wet AlAs oxidation. © 1997 American Vacuum Society.

[S0734-211X(97)08104-3]

I. INTRODUCTION

The excellent properties of AlAs oxide films formed by wet oxidation have recently attracted much attention. A simple oxidation process, usually carried out at 700–800 K on epitaxial AlAs films gives rise to robust, high quality oxide layers that have been incorporated into III–V metal–oxide–semiconductor devices,^{1–3} and used as apertures in vertical cavity lasers, confining both optical mode and drive current.^{4–7} Initial studies of the oxide growth rate (under kinetically limited conditions) have revealed a dramatic dependence on the fraction of Al in AlGaAs films, e.g., AlAs oxidizes approximately one order of magnitude faster than Al_{0.96}Ga_{0.04}As.⁷

The majority of current research has been directed at understanding the structural, compositional, and electrical properties of the fully oxidized films that result after atmospheric pressure oxidation in tube furnaces. Consequently, little is known about the evolution of the oxide at a fundamental level. In order to address this concern, we have investigated, under ultrahigh vacuum (UHV) conditions, the mechanism of water adsorption and dissociation and the subsequent growth of thin oxide films on the AlAs(001) surface using Auger electron spectroscopy (AES), temperature-programmed desorption (TPD), and low-energy electron diffraction (LEED).

II. EXPERIMENTAL METHODS

All experiments were conducted in an UHV chamber (base pressure of 1×10^{-10} Torr) that has been described in detail elsewhere.⁸ Briefly, the chamber is equipped with reverse-view LEED optics (Princeton Research), a single-pass cylindrical mirror analyzer (PHI) for AES, and a differentially pumped mass spectrometer (UTI-100C) for TPD.

The samples used in this study were 2000 Å thick AlAs (undoped) films, grown on GaAs(001) substrates by molecular beam epitaxy (MBE), which were As capped to prevent air oxidation upon exposure to the atmosphere. The AlAs(001) surface was prepared as follows: The sample was Ar⁺ sputtered (500 eV at $\sim 1 \mu\text{A}/\text{cm}^2$) to remove atmospheric contaminants and then thermally annealed to 800 K to desorb the As cap. Final annealing to 900–950 K, a temperature range in which significant As₂/As₄ desorption from the AlAs matrix occurs, resulted in a reproducible As(31 eV)/Al(64 eV) peak-to-peak ratio of approximately 0.5 using AES. Note that a persistent trace of oxygen could always be detected on the nominally clean AlAs(001) surface. Although determination of the clean surface structure (using LEED) was difficult due to the insulating nature of the undoped AlAs film at low temperatures (below 400 K), a weak (1×1) LEED pattern was observed after the above surface preparation. Only the (3×2), (1×1), and (2×3) surface reconstructions have been observed on AlAs(001):^{9,10} the (3×2) and (2×3) structures were assigned to Al- and As-

^{a)}Corresponding author; Electronic mail: mitchell@engineering.ucsb.edu

stabilized surfaces, respectively, whereas the apparent (1×1) structure, which exists at intermediary Al/As coverages in the top layer, was assigned to a disordered surface consisting of a large number of random, dimerized domains.

The extent of AlAs oxidation was monitored by AES. Derivative-mode (dN/dE) spectra were recorded after each oxygen exposure, using an electron beam energy of 3 keV. All spectra were recorded at 300–350 K, as it was not possible to obtain a reproducible signal at lower surface temperatures (presumably due to the insulating nature of the oxide). Oxygen coverage was monitored using the $O(KL_1L_{23})$ transition at 510 eV. Both the high- and low-energy As transitions at 1230 eV and 31 and 38 eV were monitored, as well as the high- and low-energy Al transitions at 1400 eV and 42 and 64 eV. The oxidic Al transitions at 39, 47, and 54 eV were also monitored (where oxidic can refer to an oxide or hydroxide or both¹¹). In all the experiments carried out here, the high-energy transitions [with electron escape depths of 20–25 Å (Ref. 14)] were found to be invariant with oxidation, suggesting that any oxide films that are formed are thin and are restricted to the surface region.

Multiplexed TPD spectra (up to eight different masses) were collected and controlled with an IBM-compatible 486 PC. The sample temperature ramps were controlled by in-house software installed on a separate PC. The TPD data were measured at constant heating rates of 1 K/s.

In order to reduce the effects of background water contamination, D_2O (Cambridge Isotope Labs, 99.9% D) was used in all experiments. The D_2O was exposed to the surface using a directional pinhole (5 μm) doser. The D_2O exposures are given in Langmuirs (L) defined in terms of the manifold pressure upstream from the doser as measured by a Baratron gauge [i.e., 1 L \equiv 1 effective monolayer = 88 Torr s (Ref. 15)].

III. RESULTS

A. AES data

The Auger spectra recorded after exposure of D_2O to the AlAs(001) surface at 100 K are presented in Fig. 1 and show a dramatic change with increasing exposure. The increase in the oxygen peak at 510 eV, see Fig. 1(b), indicates that oxygen in some form is accumulating on the surface. Recall that the AES spectra are recorded at 300 K, a temperature substantially above that for D_2O desorption on this surface (at approximately 150 K); therefore, the O signal must originate from the more stable products of D_2O dissociation, namely adsorbed hydroxyl (OD) groups and/or oxygen adatoms (indicative of partial and complete dissociation, respectively). The low-energy spectra, see Fig. 2(a), show clearly that the oxygen-based dissociation product is preferentially bound to Al (with multiple coordination) by virtue of the oxidic Al transitions at 39 and 54 eV that grow with D_2O exposure, i.e., a near-surface Al oxide/hydroxide phase has formed^{11,16} that is distinct from a chemisorbed O/OD overlayer.¹⁷ Accordingly, the Al peaks at 42 and 64 eV decrease dramatically in intensity upon conversion from Al in AlAs to “oxi-

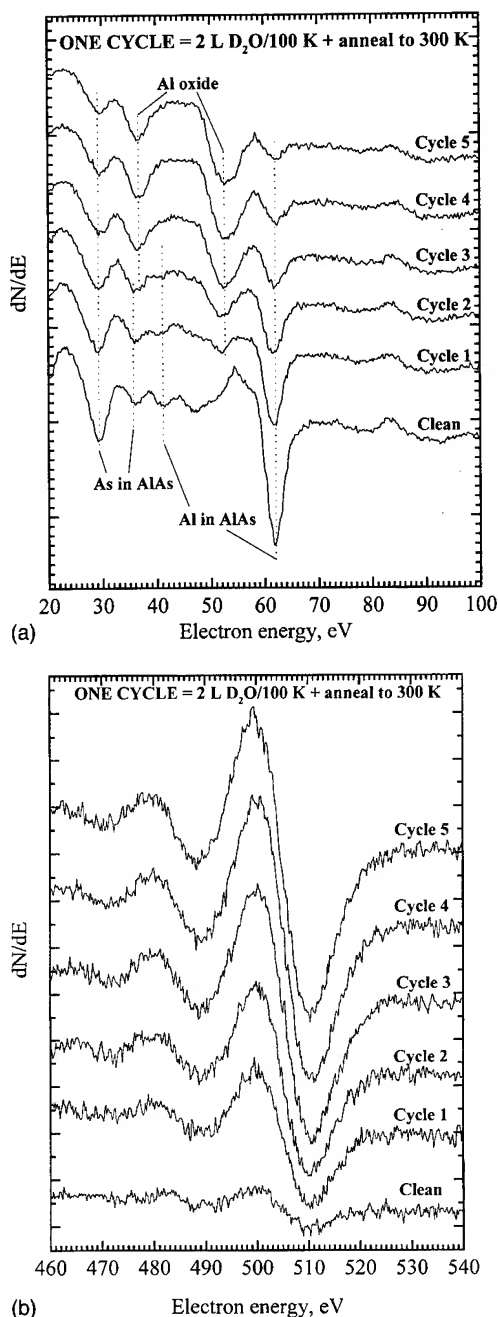


FIG. 1. Low-(a) and mid-(b) energy Auger spectra recorded after consecutive cycles of 2 L D_2O exposure to the AlAs(001) surface at 100 K followed by heating to 300 K.

dized” Al. The depletion mechanism of surface As, as indicated by the intensity decrease in the As transitions at 32 and 38 eV (the oxidic Al transition at 39 eV overlaps the As transition at 38 eV), will be made clear from the TPD data presented next.

B. TPD data

In all TPD experiments, the only desorbing species detected were D_2O ($m/e = 20$, where m/e represents the mass-to-charge ratio), AsD_3 [where AsD_2 ($m/e = 79$), representing the primary cracking fragment in the mass spectrometer, was

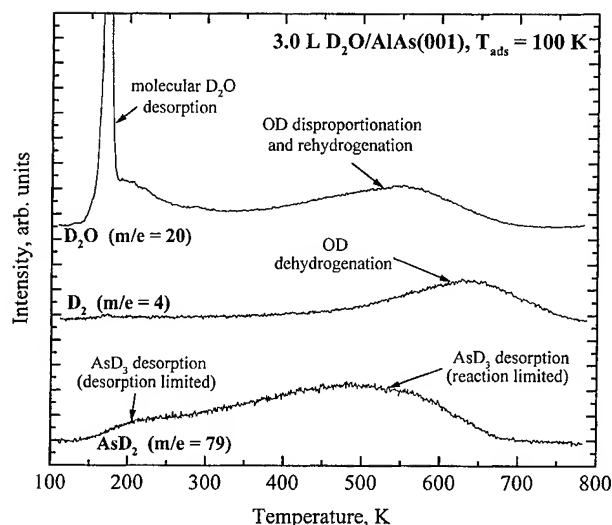


Fig. 2. D_2O ($m/e=20$), D_2 ($m/e=4$), and AsD_3 ($m/e=79$) thermal desorption spectra that result following a 3 L exposure of D_2O to the AlAs(001) surface at 100 K.

used to monitor the AsD_3 production], and D_2 ($m/e=4$). No evidence for volatile As oxides was found and, as expected, no desorbing Al oxides were detected below 800 K.¹⁸

Thermal desorption spectra recorded after exposure of the AlAs(001) surface to 3 L D_2O at 100 K are presented in Fig. 2. Clear evidence for D_2O dissociation on AlAs is provided by the broad arsine (210 and 500 K), D_2O (550 K), and D_2 (650 K) desorption peaks, respectively. The D_2 desorption peak at 650 K is assigned to the thermal decomposition of aluminum hydroxide ($Al-OD$) produced by the partial dissociation of D_2O below 200 K. The other possible sources of D_2 , namely As-D and Al-D species, are less likely since As-D will be significantly depleted below 600 K by the AsD_3 desorption peak, and high-resolution electron energy loss spectra (HREELS)¹⁹ show that little, if any, Al-D exists on the surface. The D_2O desorption peak at 550 K is assigned to the recombinative desorption of D_2O from aluminum hydroxide via either disproportionation

[$2Al-OD \rightarrow Al-O + D_2O(g) + Al$] or rehydrogenation [$Al-OD + As-D \rightarrow D_2O(g) + Al + As$]. The high-temperature desorption of AsD_3 at 500 K is a result of As-D recombination reactions in which D atoms are exchanged between As sites until AsD_3 forms, which then desorbs. After 3 L exposure, it appears that arsenic dideuteride (AsD_2) and trideuteride (AsD_3) can form directly from D_2O dissociation below 200 K: desorption of AsD_3 from these states then accounts for the low-temperature desorption peak, which is observed at 210 K. Note that this peak is not resolved after submonolayer D_2O exposure.¹⁹

The sharp D_2O desorption peak at 170 K is evidence that not all the molecularly adsorbed D_2O dissociates upon annealing, i.e., some fraction remains intact and desorbs. Significantly, similar behavior is observed after submonolayer D_2O exposures,¹⁹ namely concurrent desorption and dissociation, indicating that D_2O must adsorb molecularly at 100 K, and that competitive desorption and dissociation reaction pathways are available from this state upon annealing.

IV. DISCUSSION

The above results can be used to extract some key features in the wet oxidation of the AlAs surface. A more detailed description, utilizing an analysis of more extensive AES, TPD, and HREELS data, will be published elsewhere.¹⁹

The mechanism deduced from the data is summarized in Table I. Note that the initial adsorption step occurs with essentially unity probability at 100 K,^{20,21} implying a one-to-one correspondence between D_2O coverage and exposure. The overlap of dissociation and desorption from the molecularly adsorbed state upon annealing implies that high concentrations of water remain on the surface as dissociation proceeds. This is the main reason why we are able to observe oxidation chemistry under UHV conditions. The fact that only the dissociation reaction is observed after adsorbing submonolayer coverages of water on pure Al surfaces^{22,23} suggests that coordination of Al to As increases the activation energy to dissociation, relative to that for desorption, by

TABLE I. Summary of the wet oxidation mechanism of AlAs(001) inferred from the TPD spectra of Fig. 2.

$D_2O(g) + AlAs$	100 K →	$AlAs \cdots D_2O(a)$	Molecular adsorption
	150–180 K →	$D_2O(g)$	Molecular desorption
$AlAs \cdots D_2O(a)$	~150 K →	$Al-OD + As-D$	Partial dissociation
	~150 K →	$Al-O + As-D$	Total dissociation
$As-D$	300–650 K →	$AsD_3(g)$	Reaction-limited desorption
$As-D_x$	160–300 K →	$AsD_3(g) (x > 1)$	Desorption-limited desorption
$2 Al-OD$	350–700 K →	$Al-O + Al + D_2O(g)$	OD disproportionation
$Al-OD + As-D$	350–700 K →	$AlAs + D_2O(g)$	OD rehydrogenation
$Al-OD$	500–700 K →	$Al-O + D_2(g)$	OD dehydrogenation

a sufficient amount for both dissociation and desorption to be observed on AlAs. This coordination effect may be the basis for the sensitivity of AlGaAs oxidation rates to the fraction of Al in the film.⁷ The desorption of arsine depletes the As within the oxide film (as observed in the Auger spectra of Fig. 1) and will expose subsurface Al dangling bonds for further oxidation. Indeed, both the Auger spectra shown in Fig. 1, in which oxidic Al peaks appear and grow with increased D₂O exposure, and the TPD spectra shown in Fig. 2, in which more water dissociates than could be accommodated by surface Al atoms,²⁴ provide clear evidence for subsurface oxidation of Al. The implications of arsine desorption on subsurface Al oxidation and oxide film growth have been further investigated. Shown in Fig. 3 are (low-energy) Auger spectra recorded after sputtering (for the indicated times) the oxidized surfaces that result after repeated thermal annealing cycles between 100 and 300 K, see Fig. 3(a), and 100 and 800 K, see Fig. 3(b), with the sample positioned in front of the pinhole doser for continual exposure to D₂O during the anneal cycles. Cycling was repeated until saturation of the oxide growth was achieved [as measured by the O (510 eV) transition in AES]. The dramatic reduction in the As and Al intensities at 30 and 64 eV [with electron escape depths below 8 Å (Ref. 14)], respectively, indicate that both the low- and high-temperature oxides are at least 2–3 layers thick. The larger oxidic Al(54 eV)-to-reduced-Al(64 eV) ratio in the high-temperature oxide, however, indicates that it is denser than the low-temperature oxide. Furthermore, the persistence of the oxide peaks²⁵ for longer sputtering times in the high-temperature oxide is clear evidence for enhanced oxide growth after the 800 K annealing. The simplest explanation for this behavior is attributed to the increased AsD₃ desorption (exposing more subsurface Al for oxidation) after an 800 K anneal, relative to the desorption that results after annealing to 300 K, see Fig. 2.

V. CONCLUSIONS

We have investigated the interaction of water with the AlAs(001) surface under UHV conditions. Water exhibits a high dissociation probability on AlAs(001), even as the surface is increasingly oxidized, making it an efficient oxidant of AlAs.

The wet oxidation of AlAs(001) proceeds as follows: Water adsorbs molecularly at 100 K. A large fraction of the water dissociates with thermal annealing (the remainder desorbs) producing both aluminum hydroxide and aluminum oxide, as well as arsenic hydride, at approximately 150 K. The arsenic hydride species combine over a wide temperature range (approximately 200–600 K) to form arsine, which desorbs. The aluminum hydroxide species can either disproportionate and/or rehydrogenate between 400 and 650 K to produce desorbing water, or dissociate between 500 and 750 K to produce hydrogen, which desorbs.

The desorption of arsine, which occurs predominantly above 300 K, depletes the near-surface region of As and exposes subsurface Al dangling bonds for further oxidation. Accordingly, the AlAs oxide grown after adsorption/anneal

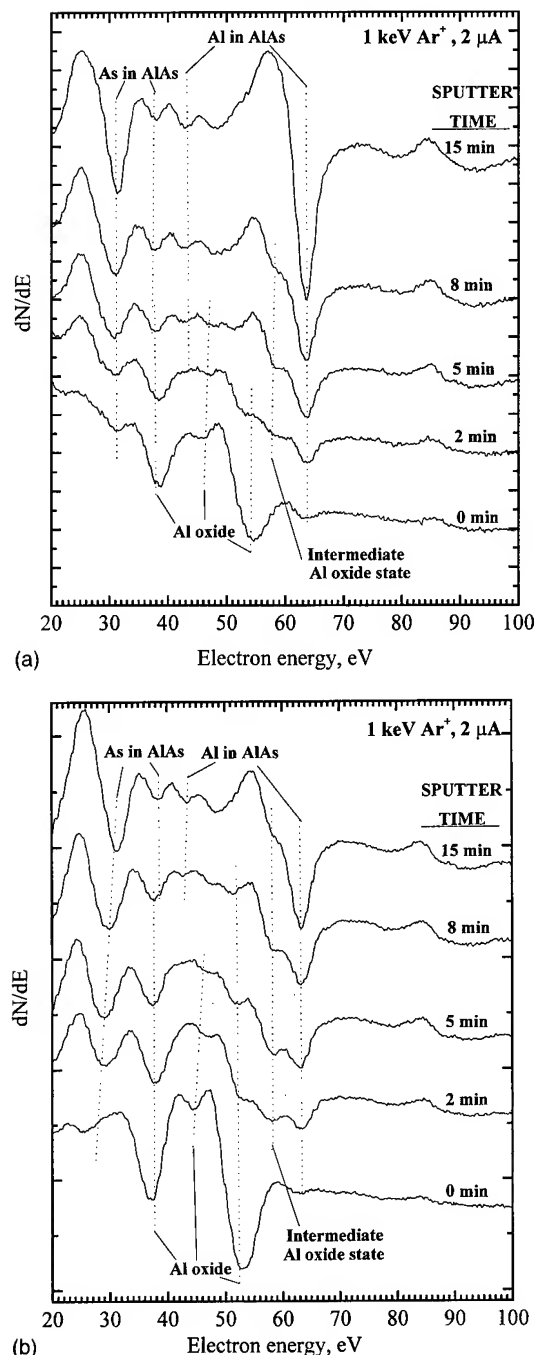


FIG. 3. Low-energy Auger spectra that result when the oxidized surfaces, prepared by continually exposing D₂O to the AlAs(001) surface (up to saturation) while thermally cycling the sample between either 100 and 300 K (a) or 100 and 800 K (b), are sputtered with Ar ions for the indicated times.

cycles between 100 and 800 K was found to be denser and thicker than the oxide grown with cycling between 100 and 300 K.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the assistance of Song Stone Shi in the MBE growth of the AlAs(001) samples used in this work. One of the authors (C.-H.C.) would like to thank the Korean Science and Engineering

Foundation for a successful postdoctoral fellowship. Primary support of this research was provided by QUEST, a National Science Foundation and Technology Center (Grant No. DMR91-20007) and the National Science Foundation (Grant No. DMR-9504400). Additional support was provided by the W. M. Keck Foundation.

- ¹W. T. Tsang, Appl. Phys. Lett. **33**, 426 (1978).
- ²E. I. Chen, N. Holonyak, and S. A. Maranowski, Appl. Phys. Lett. **66**, 2688 (1995).
- ³P. A. Parikh, S. S. Shi, J. Ibbetsen, E. L. Hu, and U. Mishra, Electron. Lett. **32**, 1724 (1996).
- ⁴A. R. Sugg, E. I. Chen, T. A. Richard, N. Holonyak, and K. C. Hsieh, J. Appl. Phys. **74**, 797 (1993).
- ⁵K. L. Lear, K. D. Choquette, R. P. Schneider, S. P. Kilcoyne, and K. M. Geib, Electron. Lett. **31**, 208 (1995).
- ⁶P. D. Floyd, B. J. Thibeault, E. R. Megblom, J. Ko, L. A. Coldren, and J. L. Merz, IEEE Photonics Technol. Lett. **8**, 590 (1996).
- ⁷K. D. Choquette, R. P. Schneider, K. L. Lear, and K. M. Geib, Electron. Lett. **30**, 2043 (1994).
- ⁸C. Huang, W. Widdra, X.-S. Wang, and W. H. Weinberg, J. Vac. Sci. Technol. A **11**, 2250 (1993).
- ⁹R. Z. Bachrach, R. S. Bauer, P. Chiaradia, and G. V. Hansson, J. Vac. Sci. Technol. **19**, 335 (1981).
- ¹⁰J. Behrend, M. Wassermeier, W. Braun, P. Krispin, and K. H. Ploog, Phys. Rev. B **53**, 9907 (1996).
- ¹¹It is very difficult to distinguish between an Al oxide and an Al hydroxide using AES (Ref. 12) [or even higher-resolution x-ray photoelectron spectroscopy (XPS) (Refs. 12 and 13)], (XPS) since the chemical shift differences in the Auger and XPS transitions between the two compounds are less than 1 eV.
- ¹²H. M. Liao, R. N. S. Sodhi, and T. W. Coyle, J. Vac. Sci. Technol. A **11**, 2681 (1993).
- ¹³S. Thomas and P. M. A. Sherwood, Anal. Chem. **64**, 2488 (1992).
- ¹⁴I. Lindau and W. E. Spicer, J. Electron Spectrosc. Relat. Phenom. **3**, 409 (1974).
- ¹⁵The doser has been calibrated by adsorbing acetylene on a cold Si(001) surface (i.e., an adsorption probability of unity).
- ¹⁶A. Hoffman, Ts. Maniv, and M. Folman, Surf. Sci. **182**, 56 (1987).
- ¹⁷P. A. Smith and S. L. Bernasek, J. Electron Spectrosc. Relat. Phenom. **49**, 149 (1989).
- ¹⁸Y. Wu, E. Garfunkel, and T. E. Madey, J. Vac. Sci. Technol. A **14**, 2554 (1996).
- ¹⁹W. J. Mitchell, C.-H. Chung, S. I. Yi, E. L. Hu, and W. H. Weinberg, Surf. Sci. (in press).
- ²⁰T. S. Wittrig, D. E. Ibbotson, and W. H. Weinberg, Surf. Sci. **102**, 506 (1981).
- ²¹S. J. Bushby, B. W. Callen, K. Griffiths, F. J. Esposto, R. S. Timsit, and P. R. Norton, Surf. Sci. Lett. **298**, L181 (1993).
- ²²J. Paul and F. M. Hoffmann, J. Phys. Chem. **90**, 5321 (1986).
- ²³U. Memmert, S. J. Bushby, and P. R. Norton, Surf. Sci. **219**, 327 (1989).
- ²⁴The fraction of D₂O that desorbs in the recombinative desorption peak at 550 K (relative to the total D₂O desorption) is approximately 0.5 after 3 L exposure. With an adsorption probability of unity at 100 K (*vide supra*), this suggests that at least 1.5 ML of the molecularly adsorbed D₂O dissociates upon annealing.
- ²⁵The oxide peak that appears at 58 eV, between the oxidic and reduced Al transitions, is assigned to an "intermediate" suboxide state of Al that is probably generated by the sputtering.

Characterization of low-temperature grown AlSb and GaSb buffer layers

K. G. Eyink

Materials Directorate, Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio 45433

M. L. Seaford

Cornell University, Ithaca, New York 14853 and Materials Directorate, Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio 45433

T. W. Haas, D. H. Tomich, W. V. Lampert, S. D. Walck, J. S. Solomon, and W. C. Mitchel

Materials Directorate, Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio 45433

L. F. Eastman

Cornell University, Ithaca, New York 14853

(Received 12 January 1997; accepted 5 May 1997)

The InGaSb/InAs strained layer superlattice (SLS) system has been proposed theoretically to be a potentially useful long-wavelength infrared (IR) detector material by Mailhot and Smith. These authors have suggested that GaSb substrates would be suitable for growth of these SLS. The benefit of InGaSb/InAs SLSs grown on GaSb(001) versus GaAs(001) substrates has been substantiated by a number of recent papers [J. L. Johnson, L. A. Samoska, A. C. Gossard, J. L. Merz, M. D. Jack, G. R. Chapman, B. A. Baumgratz, K. Kosai, and S. M. Johnson, *J. Appl. Phys.* **80**, 1116 (1996); T. D. Golding, H. D. Shih, J. T. Zborowski, W. C. Fan, C. C. Horton, P. C. Chow, B. C. Covington, A. Chi, J. M. Anthony, and H. F. Schaake, *J. Vac. Sci. Technol. B* **10**, 880 (1992)]. Unfortunately, the current quality of GaSb substrates does not even match those of InP, much less GaAs or Si. In particular, the poor quality of the GaSb substrates leads to free carriers that make the samples electrically conductive and absorb photons in the IR. Further studies of this materials system, therefore, require either improved GaSb substrates or a suitable buffer layer to electrically isolate the substrate from the SLS. In this work, we explore the use of low-temperature grown (LTG) AlSb and GaSb, with and without annealing, as an effective near lattice matched buffer layer for SLS growth GaSb and GaAs. The LTG layers were formed by solid source molecular beam epitaxy using monomeric Sb from a cracker cell. The LTG-AlSb was grown at 450–500 °C while the GaSb layers were grown at 250–350 °C. Characterization of the layers by transmission electron microscopy (TEM), high-resolution x-ray diffraction, atomic force microscopy, and Hall mobilities has been performed. These results show that LTG GaSb on GaAs substrates are highly defective and become polycrystalline after about 1500 Å of growth. TEM analysis showed the formation of precipitates in LTG AlSb films on GaSb after annealing. No Sb precipitates have been found in LTG GaSb.

© 1997 American Vacuum Society. [S0734-211X(97)11304-X]

I. INTRODUCTION

Strained layer superlattices (SLS) have grown in importance in recent years due to their potential use in many applications. The InGaSb/InAs system, for example, has been predicted by Mailhot and Smith to be useful for both infrared (IR) detector¹ and nonlinear optical² materials applications. This SLS has also found increased attention as a mid-IR laser material.³ The preferred substrate for growth of this SLS is thought to be GaSb⁴ because of the near lattice match between the SLS and the GaSb substrate. Unfortunately, although GaSb substrates are readily available in reasonable quality, they are for the most part not optically transparent in the regions of interest and they are also highly electrically conductive. These properties severely limit the ability to electrically and optically characterize materials grown on GaSb.

Several approaches exist to eliminate this problem of poor quality substrates. The substrate can be mechanically thinned so that it contributes less to the characterization techniques.

A parting layer can be grown and the SLS removed from the substrate. An electrically resistive layer can be grown underneath the SLS to isolate it from the conductive substrate. In this article, we will present results on low-temperature grown (LTG) GaSb and AlSb as possible resistive layers for this purpose. This approach is analogous to the LTG GaAs system where highly resistive layers are formed after a high-temperature anneal of material grown at low temperatures.⁵ LTG GaAs has become a very important material for buffer layer and high-speed photodetector applications.⁶ It is of interest, therefore, to determine if similar properties exist for either LTG GaSb or AlSb layers. Precipitation in this different material system should help evaluate the merit of the defect model⁷ and the buried Schottky model⁸ proposed to explain the high resistivity observed in NS arsenides.

II. EXPERIMENT

The samples in this study were grown by solid source molecular beam epitaxy in a Varian GenII system. The Sb

TABLE I. LTG GaSb and AlSb samples and characterization results. All LTG GaSb were grown at 250 °C and AlSb at 500 °C. All anneals were for 30 min at 550 °C. Layer 1 is next to the substrate.

Sample ID	Substrate	Anneal	Layer 1 GaSb	Layer 2 AlSb	Layer 3 GaSb	N_s (cm ⁻²)	μ (RT)	x-ray shift	TEM clusters	Haze	Auger depth profile
18	GaAs	N	3000LTG	0	0	6e15	30	...	N(Poly)	Y	...
19	GaAs	Y	3000LTG	0	0	2e12	250	40 arcsec	N(Poly)	Y	...
24	GaSb	N	1700	300LTG	1000	1.8e15	440	...	N	N	...
25	GaSb	Y	1700	300LTG	1000	8.7e14	440	?	Y	N	Change
29	GaSb	N	1300	1500LTG	200	2.4e16	70	...	a	N	...
30	GaSb	Y	1300	1500LTG	200	4.5e16	130	70 arcsec	a	N	No change
31	GaSb	N	1500	0	1500LTG	?	?	...	N	N	N
32	GaSb	Y	1500	0	1500LTG	?	?	0	N	N	N

^aTEM sample preparation may have caused oxidation or amorphization of AlSb layer.

source was an EPI cracker cell operated in a mode thought to produce monomeric Sb.⁹ The samples were mounted in EPI Uniblock holders without indium bonding. A sapphire backing plate was used. Temperature control is known to be very important in the growth of LTG GaAs, and reproducibility at lower temperatures is very difficult. In this work, the measurements were done using an optical pyrometer sighted through a heated quartz window. Pyrometer readings were compared for accepted oxide desorption temperatures for GaAs and GaSb substrates. Ratios of pyrometer readings to wafer heater thermocouple readings were used to extrapolate to temperatures below the range of the pyrometer. In some cases, the transition from the (1×3) reconstruction of GaSb surfaces to the (1×5), approximately 380 °C,¹⁰ was used as an intermediate data point. The V/III flux ratio measured using a beam equivalent pressure (BEP) gauge was maintained at a value of 8 for GaSb and 10 for AlSb. It should be noted that this flux ratio using monomeric Sb indicates a significantly higher flux of Sb molecules than an equivalent BEP ratio using Sb₄. With the sublimator held at a constant temperature, the V/III BEP ratio increased to approximately 40 as the cracker temperature was decreased to 600 °C to produce Sb₄ molecules. The group III fluxes used in this study produced a growth rate of 0.5 ML/s for both the GaSb and AlSb. X-ray diffraction characterization was made using a high-resolution Philips MRD system with a multiple crystal monochromator and was analyzed by dynamical diffraction simulations. Transmission electron microscopy (TEM) samples were prepared by low angle cleavage techniques to minimize sample damage.¹¹ Some of the AlSb layers, so prepared, showed the presence of surface oxidation. Auger depth profiles were done with inert gas ion bombardment using standard techniques. Hall measurements were made to determine mobilities and carrier concentrations.

III. RESULTS AND DISCUSSION

Pairs of samples were grown under the same nominal conditions. One of the samples in the pair was then annealed *in situ* under a Sb flux for ~30 min at a pyrometer temperature of 500 °C. The GaSb layers were grown at a substrate temperature ranging from 250 to 350 °C and AlSb layers were grown between 450 and 500 °C. The samples to be discussed in this work are summarized in Table I. An overall

total layer thickness was maintained at 300 nm for layers grown on both GaAs and GaSb substrates. TEM measurements made on the two LTG GaSb films grown on GaAs substrates showed an extremely defect structure. A shift of approximately 40 arcsec was observed with x-ray analysis. This shift may be due to excess Sb in the film or residual strain variation in the films. Additionally, a pair of thin LTG AlSb layers was grown between two GaSb layers grown at normal temperatures. These AlSb layers, as well as the two growths on GaAs showed a three order of magnitude decrease in sheet charge with annealing. The sheet charge was a factor of 4 lower than that achieved for a 24 μ m growth of GaSb on GaAs, which had a mobility of 700 cm²/V s and a sheet concentration of 1.3×10^{13} cm⁻².

An inspection of the data presented in Table I shows several trends. The first observation is that none of the LTG GaSb films displayed any evidence of Sb precipitates before or after annealing. This statement holds for other LTG GaSb films grown and examined but not included in Table I. Sample 25 containing LTG AlSb did display apparent Sb precipitates in the TEM micrographs following the anneal process but, sample 24 did not indicate any precipitates in its unannealed state. An example of this is shown in the micrograph in Fig. 1. Higher resolution analytical TEM is currently being performed to confirm that these inclusions are composed of Sb.

High-resolution x-ray diffraction rocking curve measurements were performed on all samples listed in Table I. None



FIG. 1. TEM micrograph of LTG AlSb layer sandwiched between GaSb layers. The left panel corresponds to the unannealed sample while the annealed sample in the right panel shows the apparent presence of Sb precipitates, which is not observed in the unannealed sample. These precipitates are held located in the central AlSb section.

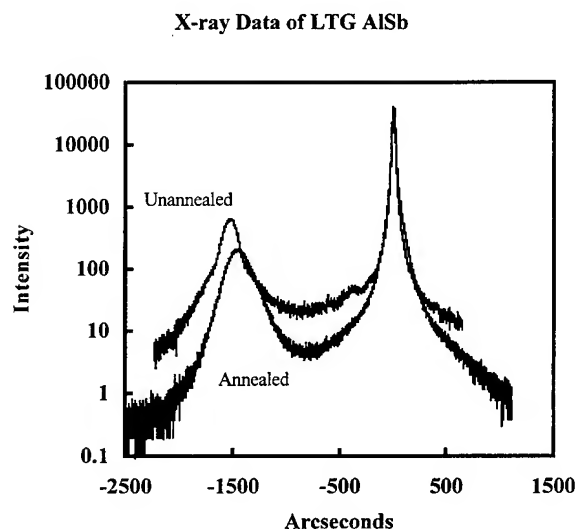


FIG. 2. X-ray rocking curves for LTG GaSb grown on GaSb substrates, Fig. 3(a), compared to annealed LTG GaAs films grown on GaAs substrates in Fig. 3(b).

of the LTG GaSb samples gave any indication of a significant change in the lattice constant between the annealed and unannealed pairs. Care must be taken in this analysis since small apparent shifts are seen. An example of this apparent shift is seen in the x-ray analysis of sample 31 in Fig. 2. These shifts are not due to excess Sb in the LTG layer but rather due to excess Sb at a buried interface. Typical rocking curves are shown in Fig. 2 for LTG GaSb (a) and LTG GaAs (b). Although these two curves appear similar, the shifts are, in fact, caused by different phenomena. In the case of the LTG GaAs, this shift is due to excess As incorporated in the film. In the case of the GaSb layer, this is due to a phase shift in the x-ray reflection at the buffer layer/substrate interface. This is evident from the fact that the fringe period corre-

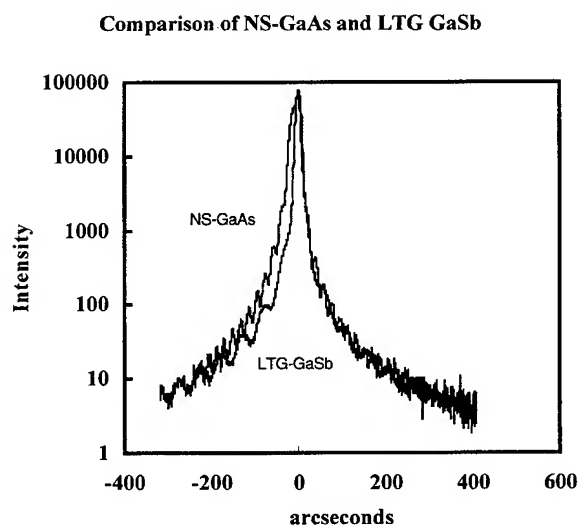
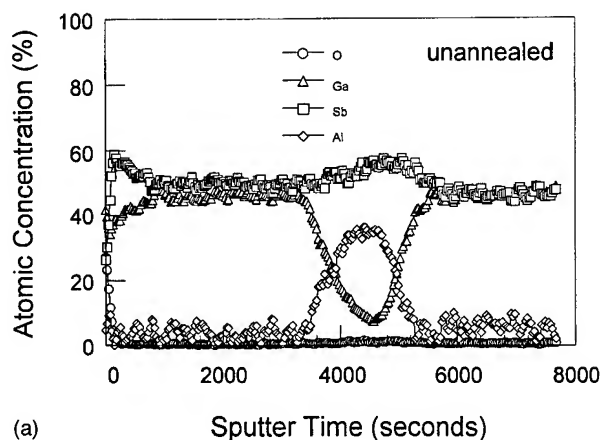
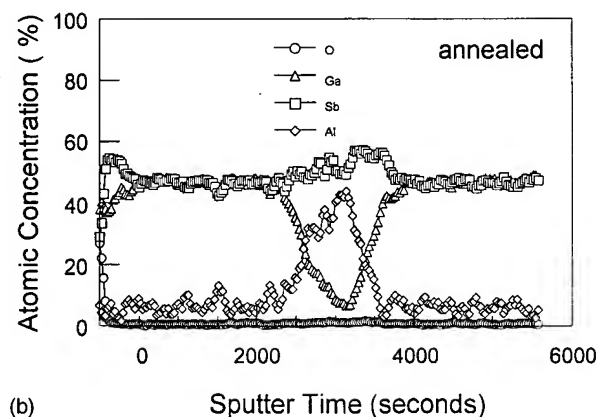


FIG. 3. X-ray rocking curves for sample number 30 from Table I. The somewhat broad peak in the left part of the scan is from the thin (approximately 150 nm thick) LTG AISb layer. The shift between the annealed and unannealed layers is clearly visible.



(a)



(b)

FIG. 4. Auger depth profiles from sample numbers 24 and 25 from Table I. The unannealed sample in (a) gives a somewhat normal depth profile for the elements present, while the profile from the annealed sample in (b) shows considerable structure in the AISb layer.

sponds to approximately 3000 Å, which is the nominal combined thicknesses of the 1500 Å buffer layer and of the 1500 Å low-temperature GaSb growth. This shift may be due to a buried Sb layer. Buried Sb layers have been observed on GaSb(111) surfaces, however, they have not currently been observed for the (001) surface.¹² Due to the possibility of interfaces causing an apparent shift in the x-ray spectrum, care must be taken in the x-ray analysis before small shifts in the x-ray spectrum can be attributed to a change in the lattice constant caused by a small amount of excess Sb.

More interesting behavior was observed in the LTG AISb samples. Sample 30 from Table I had a somewhat thicker LTG AISb layer than sample 25 and gave the rather interesting behavior indicated in Fig. 3. The annealed and unannealed samples are overlaid for easy comparison. A shift in the AISb line is clearly observed visually. Simulation of the curves shows that there is shift of 70 arcsec in the LTG layer towards a lower angle. This shift is consistent with the formation of Sb precipitates during the annealing process.

Finally, Auger electron spectroscopy depth profiles and secondary ion mass spectroscopy measurements were also carried out on several of the samples listed in Table I. As expected, the LTG GaSb samples showed the type of behav-

ior found in normally grown epitaxial GaSb. The LTG AlSb, however, displayed interesting behavior. Figure 4 shows an Auger depth profile for sample 24 discussed, as shown previously, in Figs. 1 and 3. Figure 4(a) shows the depth profile of the unannealed sample while Fig. 4(b) shows the changes resulting from the anneal process. Figure 4(a) is similar to such depth profiles from normally grown layered structures. Figure 4(b) indicates that considerable redistribution of the Al and Sb elements has occurred. Although it is somewhat speculative due to the noise in the data, one could maintain that the Al concentration is reduced in regions where the Sb concentration is increased. This, of course, would be the expected result if Sb precipitates were formed due to the anneal process. If the clusters are proven to be Sb, this material system will be instrumental in comparing the various models for the electrical and optical behavior of nonstoichiometric materials.

IV. CONCLUSIONS

Two important conclusions can be drawn from these results. First, LTG AlSb was found to exhibit a shift in lattice constant upon annealing, indicating the possible presence of excess Sb in the layer. This result was also supported by the observation of possible Sb precipitation with TEM and Auger analysis. Second, the LTG GaSb did not show Sb precipitates in this range of growth conditions in contrast to similar material such as LTG GaAs. Further, although more complete analysis is needed, none of the LTG GaSb films appeared to have the extremely high resistivity observed in annealed LTG GaAs. The use of LTG AlSb is promising in

this regard. However, recent work by Blank *et al.*¹³ suggests that the use of the AlSb buffer and intermediate layers leads to inferior electrical and optical properties in similar SLS structures. Finally, it should be noted that all this work was carried out using cracked Sb sources. There remains the possibility that tetrameric Sb might produce a different result, and work to test this hypothesis is in progress.

ACKNOWLEDGMENTS

The authors would like to express their gratitude for the excellent technical assistance of R. M. Anderson, L. Petry, R. Bach, L. Grazulis, and J. Heirholzer.

¹D. L. Smith and C. Mailhot, *J. Appl. Phys.* **62**, 2545 (1987).

²D. L. Smith and C. Mailhot, *J. Vac. Sci. Technol. A* **5**, 2060 (1987).

³R. H. Miles, D. H. Chow, Y.-H. Zhang, P. D. Brewer, and R. G. Wilson, *Appl. Phys. Lett.* **66**, 1921 (1995).

⁴C. Mailhot and D. L. Smith, *J. Vac. Sci. Technol.* **7**, 445 (1989).

⁵C. R. Wie, K. Xie, D. C. Look, K. R. Evans, and C. E. Stutz, *Mater. Res. Soc. Symp. Proc.* **198**, 383 (1990).

⁶F. W. Smith, A. R. Calawa, C.-L. Chen, M. J. Manfra, and L. J. Mahoney, *IEEE Electron Device Lett.* **EDL-9**, 77 (1988).

⁷F. W. Smith, *Mater. Res. Soc. Symp. Proc.* **241**, 3 (1992).

⁸A. C. Warren, J. M. Woodall, P. D. Kirchner, X. Yin, F. Pollack, M. R. Melloch, N. Otsuka, and K. Mahalingham, *Phys. Rev. B* **46**, 4617 (1992).

⁹EPI Application Note No. 3/95, July 1995.

¹⁰M. Yano, K. Yamamoto, T. Utatsu, and M. Inoue, *J. Vac. Sci. Technol. B* **12**, 1133 (1994).

¹¹J. P. McCaffery, *Mater. Res. Soc. Symp. Proc.* **254**, 109 (1992).

¹²W. C. Wang, J. A. Dura, J. T. Zborowski, A. Vigliante, H. C. Chen, T. D. Golding, and R. J. Meyer, *J. Vac. Sci. Technol. A* **11**, 1001 (1993).

¹³H. R. Blank, M. Thomas, K. C. Wong, and H. Kroemer, *Appl. Phys. Lett.* **69**, 2080 (1996).

Scanning tunneling microscopy and ballistic electron emission spectroscopy studies of molecular beam epitaxially grown Pt/CaF₂/Si(111) structures

V. P. LaBella^{a)} and L. J. Schowalter^{b)}

Physics Applied Physics and Astronomy Department, Rensselaer Polytechnic Institute, Troy, New York 12180

C. A. Ventrice, Jr.^{c)}

Department of Physics, University of New Orleans, New Orleans, Louisiana 70148

(Received 12 January 1997; accepted 29 April 1997)

The electronic and morphological properties of ultrathin Pt/CaF₂/Si(111) metal insulator semiconductor structures have been characterized *in situ* by scanning tunneling microscopy (STM) and ballistic electron emission microscopy (BEEM). Platinum thickness from 2 Å to 20 Å grown on 5 Å epitaxial CaF₂ have been characterized. The STM images of the Pt/CaF₂/Si(111) structures show the atomic steps of the underlying CaF₂ morphology, as well as the formation of Pt nodules. These nodules have been observed to coalesce into bigger grains when left at room temperature under ultrahigh vacuum conditions for over 24 h. For ultrathin Pt coverages (2 Å) the STM images at different tip biases reveal different topography, depending on the value of the tip bias relative to the conduction band minimum (CBM) of the CaF₂ intralayer (3.3 eV). The STM images at biases well above the CBM of the CaF₂ show similar features to bare CaF₂/Si(111), while images at the CBM show features of the deposited Pt. BEEM spectra of the 10 Å sample show a peak at ~4.5 eV due to the density of states of the CaF₂ intralayer and an additional peak at 2 eV, which has not been observed in previous studies of metal/CaF₂/Si(111) structures. © 1997 American Vacuum Society. [S0734-211X(97)11704-8]

I. INTRODUCTION

The metal-insulator-semiconductor (MIS) diode plays an important role in many modern microelectronic applications such as in charged coupled devices (CCDs) and in metal-oxide-semiconductor field effect transistors (MOSFETs). Understanding the physics of the electron transport and the growth morphology of these structures is of great importance for optimizing their electronic properties.¹ Calcium fluoride (CaF₂) is an excellent choice for an epitaxial insulating layer in silicon-based MIS structures due to its large band gap (12.1 eV) and its small lattice mismatch with Si (0.6% at 273 K). As a result of this small lattice mismatch (5.46 Å for CaF₂ and 5.43 Å for Si), low surface free energy (0.45–0.55 J/m²) of the CaF₂, and chemical bonding at the CaF₂/Si interface, very flat defect free epitaxial layers of CaF₂ can be grown on well oriented Si(111) substrates.² By utilizing these properties of CaF₂/Si heteroepitaxy, it is possible to grow MIS structures with ultrathin layers of CaF₂ as the insulating film. Extensive studies of this system indicate that the first layer of CaF₂ bonds directly to the Si to form a Si–CaF interfacial layer.^{3–6} This layer has a conduction band minimum (CBM) approximately 1 eV from the Fermi level. Successive layers of CaF₂ grow stoichiometrically, and the CBM rises to a value of 3.3 eV. For the first two monolayers of growth, CaF₂ epitaxy proceeds in a layer by layer fashion,^{7–9} although the surface morphology for thicker lay-

ers depends on the substrate orientation. Band offsets and density of state fluctuations of this CaF₂/Si(111) system have been observed by both ballistic electron emission spectroscopy (BEEM), scanning tunneling spectroscopy (STS), and photoemission studies.^{3,10}

The growth of metals on top of CaF₂ is dominated by island formation. This growth mode is a result of the low surface free energy of the CaF₂ when compared to that of the metal, the apparent lack of a chemical reaction between the Pt and CaF₂, and the large lattice mismatch between the metal and CaF₂. This island growth mode has previously been observed for Fe, Co, and Ag growth on CaF₂.¹¹ In this study platinum is used which has a surface free energy of 2.48 J/m² (Ref. 12), and a fcc lattice constant of 3.8 Å.

To investigate these novel MIS structures, scanning tunneling microscopy (STM) studies of ultrathin Pt (2–20 Å) grown onto approximately 5-Å-thick epitaxial CaF₂/Si(111) structures have been performed. These new structures, with dimensions on the order of a few angstroms, allow quantum effects such as tunneling to play an important role in how the device operates. In addition, BEEM was performed to investigate the electrical properties of these MIS structures. In BEEM, a scanning tunneling microscope (STM) tip is used to inject electrons into the metal overlayer at an energy which is determined by the tip bias.^{13,14} By collecting the fraction of injected electrons which pass through the metal-insulator overlayer into the Si substrate (i.e., the BEEM current), a direct measure of the local electron transport properties can be attained.

^{a)}Electronic mail: labelv@rpi.edu

^{b)}Electronic mail: schowalt@unix.cie.rpi.edu

^{c)}Electronic mail: cavph@uno.edu

II. EXPERIMENT

The MIS structures were grown and characterized in a commercial Si molecular beam epitaxy (MBE) system (Fisons V90H) under ultrahigh vacuum (UHV) conditions. The MBE system incorporates a reflection high-energy electron diffraction (RHEED) spectrometer to monitor growth and a custom-built STM for the BEEM measurements.¹⁵ 2 in. Si(111) wafers which are oriented within $\pm 0.25^\circ$ (arsenic doped, $0.01 \Omega \text{ cm}$) were used as substrates. The wafers were introduced into the UHV system directly from the manufacturer's container without any precleaning. The wafers were annealed at 600°C for ~ 3 h followed by a high temperature anneal at 1000°C for 30 min to remove the oxide. The wafers were cooled to 780°C before CaF₂ deposition. The RHEED measurements performed before CaF₂ deposition indicated a well ordered 7×7 surface structure. During this annealing procedure the pressure stayed below 1×10^{-9} mbar. The CaF₂ was deposited using a Knudsen cell heated to 1440°C . The intensity oscillations of the RHEED pattern were used to control the growth. This allowed for monolayer precision in the thickness of the CaF₂ layer. For this study, only a thickness of 5 \AA was used. The deposition rate for the CaF₂ was 20 s per monolayer, and the pressure stayed below 1×10^{-9} mbar. RHEED measurements performed after the deposition of the CaF₂ indicated a well ordered 1×1 surface, which is expected for epitaxial CaF₂ on Si.⁹ STM images of these surfaces indicate a layer by layer growth mode. To grow the Pt diode, the substrate was cooled to room temperature and transferred to another holder which had a shadow mask in it. The Pt deposition was carried out using an electron beam evaporator, which is located approximately 80 cm from the substrate to reduce the number of stray electrons that strike the substrate. An Inficon detector was used to monitor and control the growth of the Pt. The deposition rate of the Pt was $\sim 0.2 \text{ \AA/s}$. The pressure during Pt deposition is typically in the high 10^{-8} mbar range. The thickness measurement of both the CaF₂ and Pt layers were calibrated using Rutherford backscattering spectrometry (RBS) at the State University at Albany facility.

After growth of the MIS structures, the sample was then transferred *in situ* to the STM for measurements of the surface topography and electrical properties. All STM and BEEM measurements were performed at room temperature under these UHV conditions within 24 h of sample growth. A Burliegh inchworm was used for coarse positioning of the STM tip. Tunneling at STM biases below the CBM of CaF₂ resulted in STM tip instabilities. Therefore, all STM images in this study were taken at tip biases of -3.5 V or less. During BEEM measurements a second Burliegh inchworm was used to bring a small platinum-iridium "finger" in contact with the Pt diode. This second contact was used for grounding the metal layer during BEEM measurements only. A BeCu spring clip acted as a third contact which was used to mechanically clamp the wafer to the STM head and to allow measurements of the BEEM current.

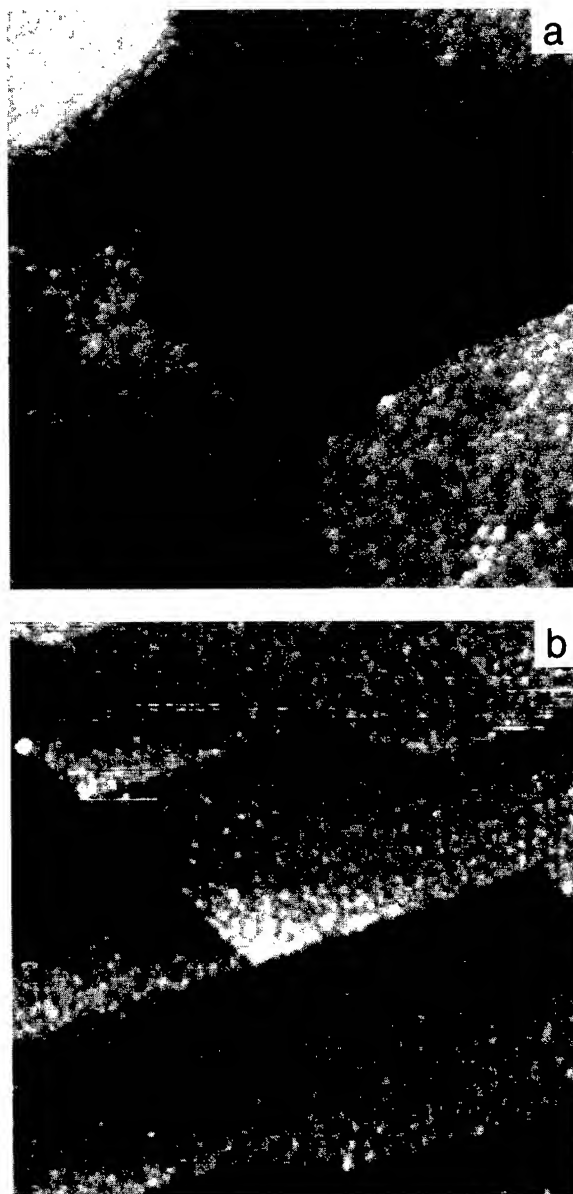


FIG. 1. (a) A $3000 \times 3000 \text{ \AA}^2$ STM image of 10 \AA of Pt on 5 \AA of CaF₂ on a Si(111) substrate ($V_{\text{tip}} = -3.5 \text{ V}$, $I_{\text{tip}} = 1 \text{ nA}$, $\Delta Z = 40 \text{ \AA}$). Notice the formation of nodules with an average width and height of 30 \AA and 6 \AA , and the underlying CaF₂ structures. (b) A $2000 \times 2000 \text{ \AA}^2$ STM image of 20 \AA of Pt on 5 \AA of CaF₂ on a Si(111) substrate, with ($V_{\text{tip}} = -3.5 \text{ V}$, $I_{\text{tip}} = 1 \text{ nA}$, $\Delta Z = 25 \text{ \AA}$). Here the Pt nodules have an average width and height of 60 \AA and 3 \AA and the underlying CaF₂ features show through.

III. RESULTS AND DISCUSSION

The STM images of 10 \AA and 20 \AA Pt/CaF₂/Si(111) surfaces are shown in Figs. 1(a) and 1(b), respectively. For both coverages Pt nodules are observed on the terraces of the underlying surface. The terraces are at different heights due to the atomic steps in the underlying CaF₂ interface. This growth mode suggests that the Pt does not wet the CaF₂ surface. For the 10-\AA -thick Pt overlayer the average nodule width and height, as measured from the STM images, is 30 \AA and 6 \AA . For the 20-\AA -thick Pt overlayer the average nodule



FIG. 2. A $3000 \text{ Å} \times 3000 \text{ Å}^2$ STM image of 10 Å of Pt on 5 Å of CaF₂ on a Si(111) substrate ($V_{\text{tip}} = -3.5 \text{ V}$, $I_{\text{tip}} = 1 \text{ nA}$, $\Delta Z = 25 \text{ Å}$). This image was acquired after the sample was left to sit at room temperature in UHV conditions for 24 h. When imaged after deposition this sample showed the formation of Pt nodules [cf. Fig. 1(a)] which have coalesced into these big grains with an average width of 500 Å . The height fluctuations on the surfaces of each grain are less than 1 Å .

width and height is 60 Å and 3 Å . This scaling effect of the nodule size with the metal overlayer thickness is similar to findings by Heim *et al.*¹¹ for Co, Fe, and Ag on CaF₂/Si(111).

The same sample shown in Fig. 1(a) is shown in Fig. 2, 24 h after deposition while allowing the sample to stay under UHV conditions and at room temperature. The nodules that were seen in Fig. 1(a) have coalesced to form bigger grains with an average width of 500 Å . The height fluctuations on the surfaces of each grain are less than 1 Å . This unusual effect was not seen by Heims *et al.*¹¹ There are several possible explanations for this effect. One is that Si atoms diffuse through the CaF₂, or through pinhole defects in the CaF₂ intralayer, and subsequently react with the Pt to form PtSi. Another explanation could be that the Pt prefers a certain thickness and is rearranging itself to achieve this. This latter effect has been observed with Ag on GaAs by Smith *et al.*¹⁶ However, at this time, we have not been able to rule out the formation of PtSi. Further experiments are being done to resolve this issue. The rest of this article will only deal with STM and BEEM results from samples immediately after Pt deposition.

Scanning tunneling microscopy images of 2 Å Pt/CaF₂/Si(111), taken at tip biases of -3.5 V and -5.0 V are shown in Figs. 3(a) and 3(b), respectively. Although both images are taken over the same region of the sample, there is a pronounced bias dependence to the imaged topography. In Fig. 3(b), a step (running from the top to the bottom of the figure) is clearly visible while the step is not visible in Fig. 3(a). In addition, Fig. 3(b) shows several of the triangular

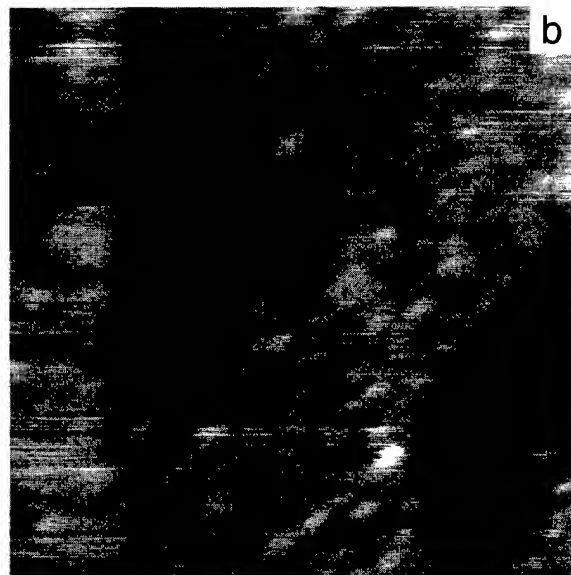
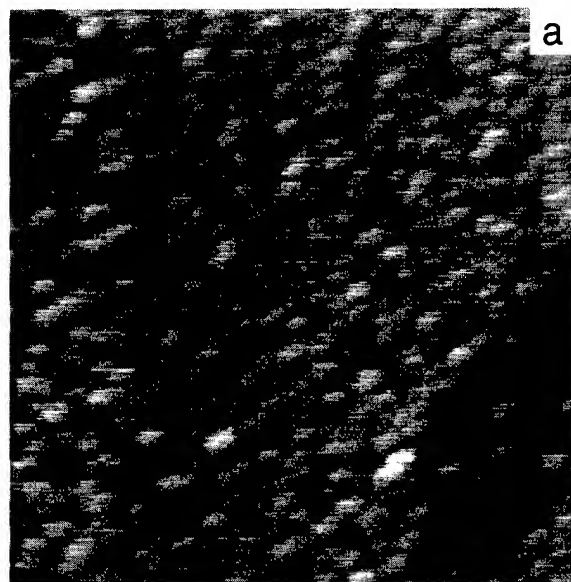


FIG. 3. (a) A $1500 \times 1500 \text{ Å}^2$ STM image of Pt/CaF₂/Si(111) ($V_{\text{tip}} = -3.5 \text{ V}$, $I_{\text{tip}} = 1 \text{ nA}$, $\Delta Z = 15 \text{ Å}$) showing the features of the Pt overlayer. (b) Same area as in (a) but with $V_{\text{tip}} = -5.0 \text{ V}$ and $I_{\text{tip}} = 1 \text{ nA}$ ($\Delta Z = 15 \text{ Å}$) showing the features of the CaF₂ intralayer. The average Pt thickness is 2 Å and the average CaF₂ thickness is 5 Å .

features associated with the thin CaF₂ layer which are not visible in Fig. 3(a) where only the Pt nodules are resolved. We believe that this unique effect results from coupling of the ballistic electrons with states either in the thin Pt layer or in the conduction band of the CaF₂. At the lower bias, the number of states available for electrons to tunnel into are greater in the Pt overlayer than in the CaF₂ intralayer. Even though the bias is just at the CBM of CaF₂, the density of states at that energy is insignificant when compared to the states available in the Pt. At -5 V tip bias, the change in topography observed in Fig. 3(b) indicates that the probability that the electrons can access a state in the conduction

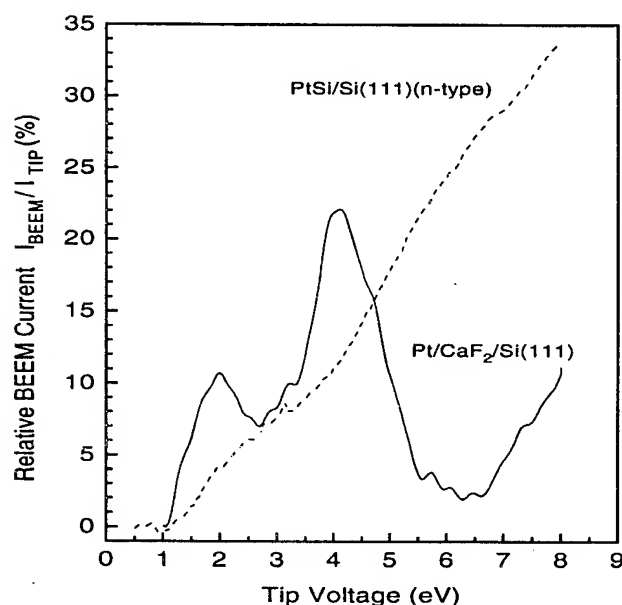


FIG. 4. Averaged ballistic electron emission microscopy spectra of a 10-Å-thick Pt layer on 5 Å of CaF₂ on a Si(111) substrate (solid line, 250 spectra) and 20 Å of PtSi on a Si(111) *n*-type substrate (dashed line, 50 spectra), both with $I_{tip} = 1$ nA. The onset of the BEEM current in the dashed spectrum at 0.9 eV is consistent with the known Schottky barrier height of PtSi on *n*-type Si. The peak in the solid line at 4.5 eV is due to the density of states of the CaF₂ intralayer. The peak at 2 eV suggests resonant tunneling in the well created by the vacuum, Pt overlayer, and CaF₂ intralayer.

band of the CaF₂ is much greater than that of the Pt. This effect allows the electrons to access the states in the CaF₂ conduction band through the ultrathin Pt overlayer. In this case the tip is not influenced by features in the Pt morphology as is seen in Fig. 3(b).

Two averaged BEEM spectra are shown in Fig. 4. The solid line is from a 10 Å Pt/CaF₂/Si(111), and the dashed line is from 20 Å PtSi/Si(111) (*n* type) sample. As seen in the figure, there are distinct peaks at 4.5 eV and 2.0 eV in the spectra from the Pt/CaF₂/Si(111) sample. The peak at 4.5 eV is due to the injection of ballistic electrons which have passed through the Pt overlayer and have accessed states in the CaF₂ conduction band. This peak occurs near the maximum of the CaF₂ density of states. This effect has been reported previously by Cuberes *et al.*¹⁰ for Au on CaF₂/Si(111). The peak at 2 eV, which was not reported by Cuberes *et al.*, may be due to resonant tunneling in the potential well created by the vacuum, Pt overlayer, and CaF₂ intralayer. Resonant tunneling would be observed near the bound state energies of the quantum well (the metal) in the double barrier structure (vacuum-metal-insulator). A simple calculation, by treating the metal as a finite well with an approximate width of 10 Å and depth of 10.9 eV, gives an energy of $\epsilon_5 \approx 2$ eV higher than the Fermi level of Pt (6.6 eV).¹⁷ The BEEM current will reflect the resonant transmission by a peak at this bound state energy. This effect will only be observed for very thin metal overlayers, since transport through thicker layers would increase the probability of

electron scattering which would lower the probability of an electron accessing this resonant state.

In interpreting this spectra, consideration of the electrical properties of the CaF-Si interface is needed since the CaF₂ is so thin (1.5 ML). Previous studies^{3,5,6} have indicated that the first layer of CaF₂ reacts to form a CaF-Si interfacial layer; this peak may result from density of state effects in this interfacial layer. The CaF-Si interface layer's CBM is near 1 eV. The BEEM spectra would then be interpreted as accessing the conduction band of the interface layer. The 2 eV peak in the BEEM spectra would correspond to the peak in the density of states for this interface layer. However, we have made recent BEEM measurements which show the disappearance of the 2 eV peak when the Pt thickness is increased while the CaF₂ thickness remains the same.

Since the onset of the BEEM spectra is near the Schottky barrier height for PtSi/Si(111) (*n* type), another possible explanation of the 2 eV peak is that the BEEM spectra were taken over a region without a CaF₂ interfacial layer. However, the BEEM spectra of PtSi/Si(111) (*n* type) show a continuous increase in the BEEM current beyond 2 eV (as shown in Fig. 4) which provides evidence against this effect. Thus, fluctuations in the Pt/CaF₂/Si(111) averaged BEEM spectra indicate that these spectra are taken where there is sufficient CaF₂ coverage, to prevent direct contact of the Pt with the underlying Si substrate. This is consistent with previous work by our group^{7,9,18} indicating that the first monolayer of CaF₂ completely covers the Si surface.

Correlating this spectra with the images from Fig. 3, the effect of tip bias on electron transmission into the semiconductor is apparent. At higher energies (5 eV), the electrons can access the CaF₂ states much more readily than at lower energies (3.5 eV). These spectroscopy results correlate well with the change in morphology that is observed at different biases as seen in Fig. 3.

IV. CONCLUSIONS

The deposition of Pt at room temperature onto epitaxial CaF₂/Si(111) exhibits several interesting effects. It is dominated by the formation of an interconnected network of Pt nodules. These nodules increase in size when the Pt thickness is increased. They also appear to coalesce when left at room temperature under UHV conditions. For ultrathin (2 Å) Pt on CaF₂/Si(111), changes in the tip bias provide a significant change in the imaged topography. This effect is due to the shape of the CaF₂ density of states and to coupling between Pt and CaF₂ electronic states. The BEES spectroscopy also shows the effects of the CaF₂ density of states and as well as a possible resonant tunneling state in the Pt well bounded by the vacuum on one side and CaF₂ intralayer on the other. Currently, more work is being carried out to examine this interesting structure.

ACKNOWLEDGMENTS

This work was partially supported in part by the Office of Naval Research and the National Science Foundation.

- ¹S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981).
- ²L. J. Schowalter and R. W. Fathauer, *J. Vac. Sci. Technol. A* **4**, 1026 (1986).
- ³Ph. Avouris and R. Wolkow, *Appl. Phys. Lett.* **55**, 1074 (1989).
- ⁴F. Gan, Y.-N. Xu, M.-Z. Huang, W. Y. Ching, and J. G. Harrison, *Phys. Rev. B* **45**, 8248 (1992).
- ⁵C. Arcangeli, S. Ossicini, and O. Bisi, *Surf. Sci.* **269/270**, 743 (1992).
- ⁶F. J. Himpsel, in *Adsorption on Ordered Surfaces of Ionic Solids and Thin Films*, edited by H.-J. Freund and E. Umbach (Springer, Berlin, 1993), pp. 3-14.
- ⁷L. J. Schowalter, B. M. Kim, T. G. Thundat, C. A. Ventrice, Jr., and V. P. LaBella, *Proc. Mater. Res. Soc. Symp.* **466** (in press).
- ⁸R. M. Tromp and M. C. Reuter, *Phys. Rev. Lett.* **73**, 110 (1994).
- ⁹B. M. Kim, Ph.D. thesis, Rensselaer Polytechnic Institute, 1996 (unpublished).
- ¹⁰M. T. Cuberes, A. Bauer, H. J. Wen, M. Prietsch, and G. Kaindl, *Appl. Phys. Lett.* **64**, 2300 (1994).
- ¹¹K. R. Heim, S. T. Coyle, G. G. Hembree, J. A. Venables, and M. R. Scheinfein, *J. Appl. Phys.* **80**, 1161 (1996).
- ¹²H. L. Skirver and N. M. Rosengaard, *Phys. Rev. B* **46**, 11,7157 (1992).
- ¹³W. J. Kaiser and L. D. Bell, *Phys. Rev. Lett.* **60**, 1406 (1988).
- ¹⁴L. D. Bell and W. J. Kaiser, *Phys. Rev. Lett.* **61**, 2368 (1988).
- ¹⁵C. A. Ventrice, Jr., V. P. LaBella, and L. J. Schowalter, *J. Vac. Sci. Technol. A* **15**, 830 (1997).
- ¹⁶A. R. Smith, K.-J. Chao, Q. Niu, and C.-K. Shih, *Science* **273**, 226 (1996).
- ¹⁷Walter A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, CA, 1980).
- ¹⁸B. M. Kim, C. A. Ventrice, Jr., T. Mercer, R. Overney, and L. J. Schowalter, *Appl. Surf. Sci.* **104/105**, 409 (1996).

Evidence of near-surface localization of excited electronic states in crystalline Si

L. Mantese^{a)} and K. A. Bell

North Carolina State University, Raleigh, North Carolina 27695-8202

U. Rossow

Institute of Physics, Technical University of Ilmenau, P. O. Box 100565, D-98684 Ilmenau, Germany

D. E. Aspnes

North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 4 April 1997; accepted 21 April 1997)

Surface- and interface-related spectra, obtained either directly by techniques such as reflectance-difference (-anisotropy) spectroscopy or indirectly by subtracting pseudodielectric function spectra obtained ellipsometrically on surfaces with different chemical termination, exhibit features related to energy derivatives of the bulk dielectric function. We argue that these spectra provide direct evidence that the excitations involved are localized both in space and time. These data unequivocally indicate that critical point energies obtained from above-band-gap ellipsometric or reflectometric optical spectra are not necessarily equal to bulk values, and that surface chemical and structural termination is at least one contributing factor. Present surface-optical calculations do not include these effects, which may explain, in part, remaining discrepancies between theory and experiment. © 1997 American Vacuum Society. [S0734-211X(97)08004-9]

I. INTRODUCTION

We recently reported surface-induced optical anisotropy (SIOA) data obtained by reflectance-difference (-anisotropy) spectroscopy (RDS/RAS) for vicinal (001) Si surfaces chemically saturated with hydrogen and oxygen.¹ These anisotropy spectra can be approximated by the first energy derivative $(-\Delta E_g + i\Delta\Gamma)d\epsilon_b/dE$ of the bulk dielectric function ϵ_b , where ΔE_g and $\Delta\Gamma$ represent differences between apparent critical point energies and broadening parameters, respectively, for light linearly polarized along the two principal axes in the plane of the surface. This dichroism in E_g and Γ clearly shows that critical point energies and broadening parameters obtained from optical spectra are not necessarily equal to their true bulk values, presumably because of the influence of the surface.

The questions then arise as to whether similar shifts are seen in other surface- or interface-related optical spectra, and how such shifts can be described in fundamental terms. By examining spectroscopic ellipsometry (SE) data for the optically isotropic (111) Si surface, we find similar shifts in E_g and Γ that depend on surface preparation and termination.² Shifts have also been reported in ellipsometric data involving interfaces,³ photorefectance,⁴ and second-harmonic generation,⁵ where they have generally been attributed to strain. Here, we present the first critical assessment of the origin of these effects. We show that these results cannot be explained by conventional stationary-state one-electron theory, since the energies and lifetimes of Bloch states are determined by their semi-infinite extent in the bulk and, hence, cannot be influenced by effects in a finite surface region. However, these results can be described if it is as-

sumed that both the initial and final electronic states involved are localized both spatially and temporally near the surface, i.e., that they can be described as wave packets consisting of a weighted superposition of Bloch states, with the driving forces for near-surface localization being the finite penetration depth of the photons and the finite lifetimes of the excitations. The energies and lifetimes of such states are determined by the overlap between bulk and surface potentials, thereby providing a justification of the observed shifts.

II. RESULTS

Representative results for two types of surface-specific optical spectra are shown in Figs. 1 and 2. The solid line in Fig. 1 represents SIOA data obtained with a RD spectrometer over the energy range of 1.5–5.5 eV. These data were taken on an oxidized (113) surface of a *n*-type Si sample with a resistivity of 1–10 Ω cm. The surface was prepared by a standard RCA cleaning procedure, but with no final HF dip so that the surface was passivated by a suboxide. Details of the surface chemical procedure⁶ and the RD equipment used to obtain these data⁷ are described elsewhere.

The solid line in Fig. 2 is taken from Ref. 2 and is the difference $\Delta\langle\epsilon_1\rangle$ between a pair of ellipsometrically measured pseudodielectric function spectra $\langle\epsilon\rangle = \langle\epsilon_1\rangle + i\langle\epsilon_2\rangle$ obtained on *n*-type (111) Si wafers with resistivities of 760–890 Ω cm. The samples were also RCA cleaned with a final 5 min immersion in NH_4F (40%) to generate a stable H-terminated surface. Again, details of the preparation and measurement procedures² and a description of the rotating-analyzer ellipsometer⁸ used to obtain these data are provided elsewhere.

The RD data are typical of results obtained on chemically saturated Si surfaces. SIOA spectra for suboxide-covered Si

^{a)}Electronic mail: lmantes@unity.ncsu.edu

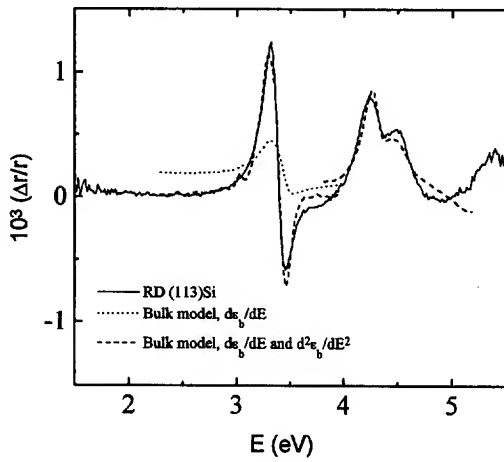


FIG. 1. Solid line: normal incidence SIOA data for (113)Si with a suboxide remaining on the surface. Dotted line: best-fit first derivative representation of the (E'_0, E_1) complex. Heavy dashed line: fits in the (E'_0, E_1) and E_2 regions using both first and second derivatives, as described in the text.

surfaces cut 4° , 6° , and 10° off (001) toward (113) differ mainly by an increase in amplitude with increasing vicinal angle,¹ which indicates that the anisotropy signal originates from steps on the surface.

III. DISCUSSION

The prominent sharp features near 3.4 and 4.5 eV in Figs. 1 and 2 occur near the energies of the (E'_0, E_1) and E_2 critical points of crystalline (c-) Si, respectively, and have characteristic energy derivative line shapes. Therefore, we model these spectra in terms of energy derivatives of ϵ_b . The general expression describing the RD data of Fig. 1 is⁹

$$\frac{\Delta r}{r} = \text{Re} \left(\frac{\tilde{r}_\alpha - \tilde{r}_\beta}{\frac{1}{2}(\tilde{r}_\alpha + \tilde{r}_\beta)} \right) \cong \text{Re} \left(\frac{-2ik_b L}{1 - 2ik_b L} \frac{(\epsilon_\alpha - \epsilon_\beta)}{n_b(\epsilon_b - \epsilon_a)} \right), \quad (1)$$

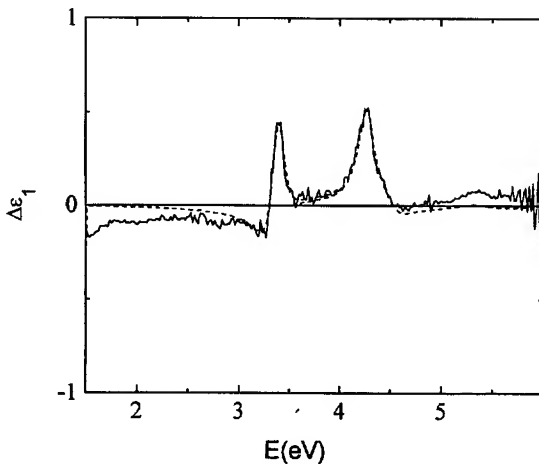


FIG. 2. Solid line: difference between pairs of ellipsometrically measured pseudodielectric function data of H-terminated (111)Si surfaces. Dashed line: fit to the data using a first derivative bulk model, as described in the text.

where ϵ_b and $\epsilon_a = 1$ are the dielectric functions of the bulk and ambient, respectively, L is the decay length of the anisotropy, which is assumed to be decreasing exponentially into the bulk, $k_b = 2\pi n_b E / hc$ is the wave vector of light in the bulk, $n_b = \sqrt{\epsilon_b}$, and \tilde{r}_α and \tilde{r}_β are the complex reflectances for light linearly polarized along the planar principal axes α and β , respectively, with ϵ_α and ϵ_β the corresponding components of the complex dielectric tensor. In the limits $d \ll \lambda$ and $d \gg \lambda$, where λ is the wavelength of light, Eq. (1) reduces to the standard equations describing an anisotropic overlayer and bulk, respectively.¹⁰ Near an isolated critical point at $E = E_g$, we suppose that $\epsilon_b(E)$ can be written

$$\epsilon_b(E) = B(E) + L(E - E_g + i\Gamma), \quad (2)$$

where $B(E)$ is a slowly varying base line and $L(E - E_g + i\Gamma)$ is the contribution of the critical point. If ϵ_α and ϵ_β differ from ϵ_b by small changes ΔE_g and $\Delta\Gamma$ in E_g and Γ , then from Eq. (2):

$$\epsilon_\alpha(E) = \epsilon_b(E) + (-\Delta E_g + i\Delta\Gamma) \frac{d\epsilon_b}{dE}. \quad (3)$$

From Eq. (3) and the equivalent expression for ϵ_β , we have

$$\begin{aligned} \frac{\Delta r}{r} &= \text{Re} \left(\frac{\tilde{r}_\alpha - \tilde{r}_\beta}{\frac{1}{2}(\tilde{r}_\alpha + \tilde{r}_\beta)} \right) \\ &\cong \text{Re} \left(\frac{-2ik_b L}{1 - 2ik_b L} \frac{(-\Delta E_g + i\Delta\Gamma)}{n_b(\epsilon_b - \epsilon_a)} \frac{d\epsilon_b}{dE} \right), \end{aligned} \quad (4)$$

where $\Delta E_g = \Delta E_g^\alpha - \Delta E_g^\beta$ and $\Delta\Gamma = \Delta\Gamma^\alpha - \Delta\Gamma^\beta$. Since ΔE_g and $\Delta\Gamma$ are expected to be different for different critical points, we divide the data into two ranges: 2.3–4.0 eV for the (E'_0, E_1) complex and 3.8–5.2 eV for the E_2 transitions, and use least-squares regression to find the values of ΔE_g and $\Delta\Gamma$ that minimize the mean-square deviation between Eq. (4) and the data.

For reasons that will be discussed below, comparisons were made using the bulk model in the limit $L \gg \lambda$ of Eq. (1). The result for the (E'_0, E_1) complex using the first derivative model is shown as the dotted line in Fig. 1. The representation is clearly not adequate, in part because the calculated line shape is too broad. Accordingly, we examined the possibility that the (E'_0, E_1) structure could be fit with a second energy derivative $d^2\epsilon_b/dE^2$. The best-fit results, including both first and second derivative line shapes in the model calculation for both (E'_0, E_1) and E_2 are shown in Fig. 1 as the heavy dashed line. The specific form of Eq. (4) used here is

$$\begin{aligned} \frac{\Delta r}{r} &= \text{Re} \left\{ \frac{1}{n_b(\epsilon_b - \epsilon_a)} \left[(-\Delta E_{g1} + i\Delta\Gamma_1) \frac{d\epsilon_b}{dE} \right. \right. \\ &\quad \left. \left. + (\Delta E_{g2}^2 + i\Delta\Gamma_2^2) \frac{d^2\epsilon_b}{dE^2} \right] \right\}, \end{aligned} \quad (5)$$

where ΔE_{g2}^2 and $\Delta\Gamma_2^2$ are the adjustable parameters of the second derivative contribution. The agreement is now satisfactory in both energy ranges with first derivative contributions $\Delta E_{g1} = (-0.037 \pm 0.05) \text{ meV}$ and $\Delta\Gamma_1 = (-0.41$

± 0.11) meV and second derivative contributions $\Delta E_{g2}^2 = (8.4 \pm 9) \times 10^{-6}$ meV² and $\Delta \Gamma_2^2 = (166 \pm 6) \times 10^{-6}$ meV² from 2.3 to 4.0 eV. From 3.8 to 5.2 eV, the data were fit with the first derivative coefficients $\Delta E_{g1} = (-0.73 \pm 0.16)$ meV and $\Delta \Gamma_1 = (-1.27 \pm 0.14)$ meV and second derivative coefficients $\Delta E_{g2}^2 = (21.1 \pm 51) \times 10^{-6}$ meV² and $\Delta \Gamma_2^2 = (50.9 \pm 11) \times 10^{-6}$ meV² from 3.8 to 5.2 eV. The second derivative line shape empirically accounts for more than 75% of the (E'_0, E_1) structure but only about 40% of the E_2 structure, showing that (E'_0, E_1) is dominated by second energy derivative contributions and E_2 by the first energy derivative. Slightly better fits were obtained using the bulk model for (E'_0, E_1) and the surface model for E_2 , but with the exception of the adjustable parameter values, the differences were not significant. Over a small energy range, the difference between the two is only a phase factor. Differences between the use of the joint-density-of-states operator, $E^2(d/dE) \times (E^2 \cdot \epsilon_b)$, and $d\epsilon_b/dE$ were also insignificant.

The second derivative line shape of the (E'_0, E_1) structure is surprising since the only mechanism that has been shown to give rise to a $d^2\epsilon_b/dE^2$ line shape is an electroreflectance effect in heavily doped semiconductors,¹¹ which is clearly not relevant here. Instead, we interpret this result as a consequence of the near degeneracy of the E'_0 and E_1 critical points of Si as follows. We suppose that the two critical points are characterized by amplitudes, critical point energies, and broadening parameters c_i , E'_{gi} , and Γ'_i , respectively, where $i=1$ and 2 and $E'_{g1} \leq E'_{g2}$ and where the line-shape $L(E - E'_i + i\Gamma'_i)$ is the same for each, except possibly for a complex prefactor. Let the energies E'_{g0} and Γ'_0 be such that $E'_{g1} \leq E'_{g0} \leq E'_{g2}$ and $\Gamma'_1 \leq \Gamma'_0 \leq \Gamma'_2$. Then

$$\epsilon_b(E) = B(E) + c_1 L(E - E'_{g0} - \Delta E'_{g1} + i\Gamma'_0 + i\Delta\Gamma'_1) + c_2 L(E - E'_{g0} - \Delta E'_{g2} + i\Gamma'_0 + i\Delta\Gamma'_2), \quad (6a)$$

where $\Delta E'_{gi} = E'_{gi} - E'_{g0}$ and $\Delta\Gamma'_i = \Gamma'_i - \Gamma'_0$. Expanding Eq. (6a) to second order, we have

$$\begin{aligned} \epsilon_b(E) = & B(E) + (c_1 + c_2) L(E - E'_{g0} + i\Gamma'_0) \\ & + [(-c_1 \Delta E'_{g1} - c_2 \Delta E'_{g2}) + i(c_1 \Delta\Gamma'_1 + c_2 \Delta\Gamma'_2)] \\ & \times \frac{\partial L}{\partial E} + [\frac{1}{2}c_1 (-\Delta E'_{g1} + i\Delta\Gamma'_1)^2 + \frac{1}{2}c_2 (-\Delta E'_{g2} \\ & + i\Delta\Gamma'_2)^2] \frac{\partial^2 L}{\partial E^2}. \end{aligned} \quad (6b)$$

Without loss of generality, we pick E'_{g0} and Γ'_0 so that the term proportional to $\partial L/\partial E$ in Eq. (6b) vanishes, thereby relating energy derivatives of $L(E - E'_{g0} + i\Gamma'_0)$ to the corresponding energy derivatives of $\epsilon_b(E)$ to within a second-order correction.

Now suppose that E'_{g1} , E'_{g2} , Γ'_1 , and Γ'_2 change slightly as a result, e.g., of surface termination. Neither first nor second derivative terms will now vanish, leading to Eq. (5). The coefficients of Eq. (5) can be written in terms of coefficients defined in Eq. (6b), although for the second derivative term the connection is not simple. This model clearly requires at

least two critical points being nearly degenerate in energy, and hence, should be more important for the (E'_0, E_1) complex than for the E_2 transition of Si as suggested by the data.

The SE data of Fig. 2 are in pseudodielectric function form and, hence, can be modeled directly in terms of $(-\Delta E_g + i\Delta\Gamma)d\epsilon_b/dE$. The dashed line shows a fit to the data taken from Bell *et al.*² with $\Delta E_g = (2.5 \pm 0.2)$ meV and $\Delta\Gamma = (-1.0 \pm 0.3)$ meV from 1.5 to 4.0 eV and $\Delta E_g = (-1.3 \pm 0.7)$ meV and $\Delta\Gamma = (-0.1 \pm 0.3)$ meV from 3.6 to 6.0 eV. The model also includes transparent overlayer and surface local field contributions as described in Ref. 2. However, the main contribution here is from $d\epsilon_b/dE$. Second derivative contributions are found in difference SE spectra involving surfaces oxidized by H_2O_2 .²

We now consider how shifts in critical point energies can arise. When different samples are involved, two obvious possibilities are differences in strain and temperature. Using a typical value of around 10 eV for the hydrostatic deformation potential,¹² a shift of 3 meV would require a bulk strain of 0.03%, an amount that is unlikely to be caused by chemical preparation. However, we cannot rule out a chemically induced shift of atomic positions within the unit cell or a net volume change normal to the surface for several layers near the surface. Temperature effects can be eliminated since the known temperature coefficient of the E_1 transition, 0.2 meV/°C,¹³ would require a change of 15 °C to generate a 3 meV shift between samples. To confirm that temperature effects are negligible in the SE data, we monitored a test sample *in situ* with a thermocouple and found that, although temperature reductions of up to 5 °C could occur due to evaporative cooling from a MeOH rinse, temperatures returned to room values in less than 1 min. Further, temperature effects cannot contribute to the SIOA data since these are obtained via different incident polarization states on the same sample, which for these measurements, is at room temperature. In addition, in no case is strain or temperature expected to influence $\Delta\Gamma$ significantly.

If an overlayer is present, $\langle\epsilon\rangle$ spectra will exhibit E_g and Γ values that are different from those obtained from $\langle\epsilon\rangle$ data for samples where overlayers are negligibly thin or absent. However, false-data calculations show that 5 Å of SiO_2 are needed to cause a 1 meV shift in the apparent energy of the (E'_0, E_1) critical point. A layer this thick would reduce the height of the E_2 peak in $\langle\epsilon_2\rangle$ from 48.3 to 45.3, which is not observed. Apparent shifts in critical point energies have also been interpreted in terms of spatial dispersion enhanced by the nearly parallel bands along Λ .¹⁴ While this model yields qualitative agreement for the E_1 structures of GaAs and InP, it is not obvious that the same mechanism would apply to the E_2 transition. In principle, apparent shifts in energy can also result from carrier confinement resulting from band bending, where the surface and bulk Fermi levels are at different positions with respect to the valence-band edge. However, confinement-induced shifts require doping levels orders of magnitude larger than are relevant here. In addition, the Franz-Keldysh effect, a second consequence of band bending, leads to third derivative line shapes, unlike those seen

here. We further note that, although the image-charge potential for *c*-Si has a magnitude of 3 meV for a point charge $\pm e$ at a depth of 100 Å below the surface, this shift is independent of the sign of the charge. Since both electron and hole experience the same shift, no net effect would be observed.

Accordingly, we consider a nontraditional mechanism, spatial and temporal localization, which we recently proposed in Ref. 9 and which, to our knowledge, has not been previously used in calculations of the dielectric response. We show first that localization effects are relevant by using an exactly solvable stationary-state model to prove that surface-induced shifts ΔE_g in apparent critical point energies are impossible for semi-infinite stationary (e.g., Bloch) states in an infinite crystal. We approximate the bulk states as plane waves and the surface potential as a delta function, $(V \cdot d)\delta(z)$, where d can be considered the width of the actual surface potential. We now place the surface potential in an infinitely deep well of width $2b$, representing two crystals joined at $z=0$. For wave functions of even symmetry, the δ function causes a discontinuity of the slope at the center of the well whereas the odd functions are unaffected. For small d/b , the energies of the even solutions can be determined analytically, and we find $E_n = n^2 \pi^2 \hbar^2 / 2mb^2 + V \cdot d/b$, where n is an integer and m is the free-electron mass. Thus, as $b \rightarrow \infty$ the effect of the surface potential vanishes. Since the energies of semi-infinite stationary states cannot be affected by a localized surface potential, we conclude that the observation of a finite ΔE_g is direct evidence of localization of the wave functions themselves.

We next consider how localization can occur. The unperturbed Hamiltonian contains no mechanism for localization, so it must result from either a strongly binding (surface) potential creating a state that does not overlap the continuum or the source term that created the packet. We eliminate the first since the line shapes show that the states involved are clearly dominated by the bulk potential. In principle, the surface potential could force the relative phases of nominally independent Bloch functions to be correlated such that localization occurs, however, the model calculation above shows that this does not necessarily occur. Accordingly, we consider the source, i.e., photon term. For an absorption coefficient $\alpha \neq 0$, this can be written in the long-wavelength limit as

$$H' = -e\mathbf{E}_0 \cdot \mathbf{r} e^{-\kappa|z| - i\omega t}, \quad (7)$$

where $\kappa = \alpha/2$ is the extinction coefficient. We assume that H' is turned on at $t=0$ and off at some later time, although in practice the lifetime is determined not by the photon but by the lifetime of the excitation. The initial and final wave functions ψ of the perturbed system can be expanded in one-electron Bloch states, $\phi_n(\mathbf{k}, \mathbf{r})$, as

$$\psi_n(\mathbf{k}_0, \mathbf{r}, t) = \int d^3k A_n(\mathbf{k} - \mathbf{k}_0, t) \phi_n(\mathbf{k}, \mathbf{r}) e^{-(i\hbar)E_n(\mathbf{k})t}, \quad (8)$$

where the coefficients $A_n(\mathbf{k} - \mathbf{k}_0, t)$, centered about \mathbf{k}_0 , are found by standard time-dependent perturbation theory and

n is the band index. The result is a set of basis functions correlated in space and time. In standard calculations of the polarizability, each Bloch function is treated independently, with the result that these correlations and the accompanying localization effects are discarded. Since the photon potential $e^{-\kappa|z|}$ is not qualitatively different from a δ function for a semi-infinite crystal, the exact stationary-state solution shows that spatial localization of a single Bloch function by itself is not adequate and that space and time correlations are also required.

Calculations that treat spatial and temporal aspects exactly would be formidable, since they must involve the manifold of states into which the excited states can scatter. However, the above arguments suggest that such calculations are essential for quantifying the observed shifts ΔE_g in E_g . We note that excitation should occur only where the photon potential is nonzero, since the probability amplitudes of the excited states should follow the photon potential. Also, since the excited electron must overlap spatially with the created hole, the initial states must also be localized. This places constraints on possible values of $A_n(\mathbf{k}, t)$ and, in turn, implies that the mean energies of the initial and final packets are a weighted average of bulk and surface potentials. Such localization probably involves a manifold of states that does not conserve energy during the lifetime of the excitation, although energy is conserved overall when the destination states are considered. Finally, because such localization would occur on the scale of the penetration depth of light, ~ 50 – 500 Å in the present case, the use of the bulk model, Eq. (5), rather than the surface model, is justified.

We now consider shifts $\Delta\Gamma$ in the broadening parameter Γ , which describe changes in the lifetimes of the excited states. Lifetimes are affected by the overlap of these states with the surface potential since this establishes, in part, the probability of scattering these states into states of lower energy. In support of this interpretation, we note that the minimum Γ values for Si are obtained with H-terminated (111) Si samples,² consistent with the fact that this is the highest quality surface that can be prepared on this material. Since the surface potential does not depend on the polarization of the incident light, the SIOA dependence of both ΔE_g and $\Delta\Gamma$ on polarization is clearly connected to the fact that, according to polarization selection rules, states excited by different polarizations will have different overlaps with the surface potential.

In this discussion we have neglected several issues. First, our localization arguments are general and would appear to apply to optical spectra of any crystalline semiconductor. However, RD data for clean reconstructed Si surfaces, which nominally contain a large density of unsaturated dangling bonds, are characterized by broad features at energies other than those of bulk critical points with no apparent derivative contribution. It would be expected, however, that the potential associated with a chemically saturated surface would be stronger than that associated with dangling bonds. Second, the surface local field effect, which dominates the RD spectra of (110) surfaces of Si and Ge,¹⁰ needs to be incorporated. These issues will be discussed in future publications.

IV. SUMMARY AND CONCLUSIONS

We have shown that surface-related optical spectra for (111) and vicinal (001)Si surfaces with saturated dangling bonds result in apparent critical point energies and broadening parameters that are different according to surface termination and incident polarization, manifesting themselves as spectra with line shapes related to energy derivatives of the bulk dielectric function. We have also shown that such shifts are inconsistent with stationary solutions of the Hamiltonian for a semi-infinite crystal modeled as a simple quantum well, and hence, require some form of spatial and temporal localization.

Our results are significant for several reasons. First, such localization has not been previously identified in solids, although it is implicit in the three-step model of photoemission¹⁵ and occurs in atoms with high Rydberg orbits through the localization of the initial states.¹⁶ Second, the values of critical point energies and broadening parameters determined from optical data may, and most probably will, differ from true bulk values. Thus, when analyzing optical data involving interfaces and/or extremely thin films, care must be taken to use reference spectra obtained from samples with the same surface termination, otherwise errors in layer thicknesses and estimates of interfacial strain can be expected. Finally, these effects must be considered in theoretical calculations of SIOA spectra if the full diagnostic power of surface-optical techniques is to be exploited.

ACKNOWLEDGMENTS

This work is supported by the Office of Naval Research (ONR) under Contract No. N-00014-93-1-0255 and the Alexander von Humboldt Foundation.

- ¹L. Mantese, U. Rossow, and D. E. Aspnes, *Appl. Surf. Sci.* **107**, 35 (1996).
- ²K. A. Bell, L. Mantese, U. Rossow, and D. E. Aspnes, *J. Vac. Sci. Technol. B* **15**, 1205 (1997).
- ³N. V. Nguyen, D. Chandler-Horowitz, P. M. Amirtharai, and J. G. Pellegrino, *Appl. Phys. Lett.* **64**, 2688 (1994).
- ⁴J. T. Fitch, C. H. Bjorkman, G. Lucovsky, F. H. Pollak, and X. Yin, *J. Vac. Sci. Technol. B* **7**, 775 (1989).
- ⁵W. Daum, H.-J. Krause, U. Reichel, and H. Ibach, *Phys. Rev. Lett.* **71**, 1234 (1993).
- ⁶T. Yasuda, D. E. Aspnes, D. R. Lee, C. H. Bjorkman, and G. Lucovsky, *J. Vac. Sci. Technol. A* **12**, 1152 (1994).
- ⁷D. E. Aspnes, J. P. Harbison, A. A. Studna, and L. T. Florez, *J. Vac. Sci. Technol. A* **6**, 1327 (1988).
- ⁸See for example, D. E. Aspnes and A. A. Studna, *Appl. Opt.* **14**, 220 (1975), and references therein.
- ⁹U. Rossow, L. Mantese, and D. E. Aspnes, in *Proceedings of the 23rd International Conference on the Physics of Semiconductors*, edited by M. Scheffler and R. Zimmerman (World Scientific, Singapore, 1996), p. 831.
- ¹⁰D. E. Aspnes, *J. Vac. Sci. Technol. B* **3**, 1498 (1985).
- ¹¹T. Lukes and K. T. S. Somaratna, *J. Phys. C* **3**, 2044 (1970).
- ¹²P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 1996), p. 118.
- ¹³P. B. Allen and M. Cardona, *Phys. Rev. B* **27**, 4760 (1983).
- ¹⁴D. E. Aspnes and A. A. Studna, *J. Vac. Sci. Technol. A* **5**, 546 (1987).
- ¹⁵The three-step model of photoemission is described in P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 1996), p. 428.
- ¹⁶See for example, X. Wang and W. E. Cooke, *Phys. Rev. Lett.* **67**, 976 (1991); L. D. Noordam, H. Stapelfeldt, D. I. Duncan, and T. F. Gallagher, *ibid.* **68**, 1496 (1992).

Microwave modulated photoluminescence as a contactless probe of interface states

C. E. Inglefield, M. C. DeLong, and P. C. Taylor
Physics Department, University of Utah, Salt Lake City, Utah 84112

J. F. Geisz and J. M. Olson
National Renewable Energy Laboratory, Golden, Colorado 80401

(Received 13 January 1997; accepted 22 April 1997)

Microwave modulated photoluminescence (MMPL) is a developing spectroscopy in which the sample is subjected to continuous optical pumping and chopped microwave electric fields. The signal detected in an MMPL experiment is the change in the photoluminescence spectrum of the sample due to the presence of microwave electric fields, which increase the kinetic energy of the free carriers. In order to investigate the effects of interfaces on this measured quantity, two types of GaAs samples have been compared. The first type of sample was grown such that the GaAs epilayers are exposed, while in the second type the GaAs is "capped" by a layer of higher band gap material. Several pairs of such samples have been studied. The MMPL results are consistent with the following mechanism: an increase in the kinetic energy of the free carriers that results from the presence of the microwave fields allows more free carriers to reach the interface or surface of the GaAs layer before recombining. The presence of a greater number of nonradiative recombination paths in the samples with a bare GaAs surface than at the GaAs/capping-layer interface is therefore seen experimentally as an increase in the photoluminescence quenching by the microwave electric fields. The potential usefulness of MMPL as a probe of non-radiative recombination and as an indicator of interface quality is illustrated by a quantitative estimate of surface/interface non-radiative recombination. © 1997 American Vacuum Society. [S0734-211X(97)09504-8]

I. INTRODUCTION

In a microwave modulated photoluminescence (MMPL) experiment, the measured quantity is the change in the photoluminescence (PL) spectrum of a sample due to the presence of microwave electric fields. This is done with standard lock-in techniques used in modulation spectroscopy. In our apparatus, the sample is placed at the electric field maximum of a TE_{011} microwave cavity with optical access. The sample is subjected to continuous optical pumping from an argon ion laser (5145 Å) and chopped microwave electric fields from a 400 mW supply with magnitude controlled by a variable attenuator. For all the data discussed here, a microwave power of ~ 4 mW, corresponding to microwave electric fields ~ 10 V/cm, was used.¹ The luminescence from the sample is dispersed with a 0.85 m double grating spectrometer and detected with a cooled GaAs photomultiplier. For the PL spectra the laser light was chopped mechanically and the signal was detected with standard lock-in techniques. For MMPL spectra the light was not chopped and the lock-in amplifier detected only changes due to the chopped microwaves. All experiments were performed in a liquid helium cryostat at a temperature of ~ 5 K.

The microwave electric fields change the PL spectrum of the sample because they accelerate the free carriers in the material to a higher kinetic energy than that corresponding to the lattice temperature. This increase in carrier energy can affect the PL spectrum in a variety of ways. One possibility is that a change in the PL spectrum results from the microwave accelerated carriers equilibrating with the crystal lattice

and heating the sample. In this case the MMPL spectrum is a manifestation of the temperature dependence of the PL spectrum.^{2,3} We have previously established that effects related to sample heating are easily distinguished in our apparatus due to long transients in the signals which result from inefficient cooling of the samples.^{2,3} The microwaves can also change the PL spectrum independent of heating of the crystal lattice. Possible mechanisms for such an effect are a dependence of the capture cross section of a luminescence process on the energy of the carriers involved as well as impact ionization from weakly localized states by the accelerated carriers.³⁻⁵

In this study, we used sets of GaAs samples which had similar bulk properties but were in different epitaxial structures. The main comparison will be made between samples which were grown to have a bare (exposed to the atmosphere) surface and samples which were "capped" by a layer of higher band gap material. By design, the GaAs epilayers in the samples were nominally identical, and the GaAs/GaInP interface at the capping layer was similar to interfaces known to have low recombination velocities from a previous study.⁶ Since the samples have similar bulk properties (both by design and as measured by low temperature PL), the differences in the MMPL spectra of the samples are expected to be related to the different GaAs interfaces. MMPL has advantages over other tools because it is contactless, requires minimum sample preparation, and can in principle probe any layer in a heterostructure, including a buried layer, that luminesces or affects the luminescence in some quantifiable way. Proximal probe techniques are limited to

exposed surfaces and may require careful preparation and handling of the samples.⁷ Time resolved luminescence measured as a function of epilayer thickness is a standard technique for measuring recombination velocities but requires multiple samples be grown while altering only the epilayer thickness.⁶

Microwave electric fields can cause qualitative changes in the intensity within a PL spectrum by redistributing the carriers among the various ongoing luminescence processes. The microwaves can also cause a quantitative change (i.e., a change in the integrated intensity of the luminescence) if, as in any real material, nonradiative recombination is present. The lifetime for carriers in the presence of competing radiative and nonradiative recombination is usually expressed in terms of the lifetimes of the competing processes:

$$1/\tau = 1/\tau_r + 1/\tau_{nr}, \quad (1)$$

where τ , τ_r , and τ_{nr} are the carrier lifetime, radiative lifetime and nonradiative lifetime, respectively. The integrated luminescence intensity is proportional to the internal quantum efficiency of the sample, η

$$\eta = \frac{\tau_{nr}}{\tau_{nr} + \tau_r}. \quad (2)$$

Equation (1) is somewhat misleading in the case where the majority of the nonradiative recombination is at an interface or surface of the sample, as is the case in the samples with a bare GaAs surface, because it implies that the radiative and nonradiative recombination take place in parallel. In reality, one might see strong luminescence even if the nonradiative lifetime (τ_{nr}) at the surface is very short compared to the radiative lifetime (τ_r) in the bulk. This is because the carriers generated in the bulk have to diffuse to the surface before recombining there and the timescale of this diffusion is not necessarily short compared to the radiative lifetime in the bulk. The problem of interface or surface recombination competing with bulk recombination to determine the lifetime of minority carriers in semiconductors has been extensively studied for the case of double heterostructures.⁶ It has been shown that the lifetime for recombination at the surface for a wide range of parameters may be approximated by^{6,8}

$$\tau_{nr} = \tau_s = \frac{d^2}{\pi^2 D} + \frac{d}{2S}, \quad (3)$$

where d is the thickness of the active layer of the heterostructure, D is the minority carrier diffusion constant, and S is the surface/interface recombination velocity. Because of the form of Eq. (3), it is possible to measure (with time resolved photoluminescence) S and the bulk minority carrier lifetime using heterostructures with different active layer thicknesses.⁶

In our experiment, we change the relative strengths of the competing radiative and nonradiative recombination by raising the free carrier energy. The unique aspect of the MMPL is that we are changing the nonradiative lifetime in Eq. (3) by modulating D . In some cases the integrated PL intensity is quenched by more than an order of magnitude when a

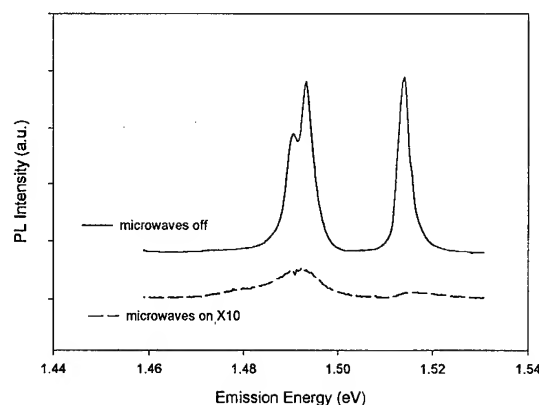


Fig. 1. Conventional PL from bulk GaAs in the absence (upper trace) and presence (lower trace) of continuous microwave electric fields. The data for the latter has been multiplied by ten to facilitate comparison. An offset is added to the upper spectrum for clarity. The microwave electric field quenches the luminescence by more than one order of magnitude.

sample is subjected to a microwave electric field. In other samples the integrated PL intensity does not change measurably. As will be shown, this substantial difference can be attributed to differences in the surfaces and interfaces involved. To get the large degree of luminescence quenching, both terms on the right hand side (RHS) of Eq. (3) must be small (with the microwaves "on") compared to the competing radiative lifetime. The term in Eq. (3) which contains D and the radiative lifetimes are bulk properties and are not expected to change significantly from sample to sample in this study. Differences in the MMPL spectra from sample to sample will therefore be attributed to differences in the recombination velocity, S .

II. RESULTS

The very large changes in a photoluminescence spectrum achievable with modest microwave powers are demonstrated by observing the effect of a continuously applied microwave electric field on the photoluminescence of a high purity GaAs sample, in this case grown by molecular beam epitaxy. These changes are demonstrated in Fig. 1. The highest energy peak in the PL from the sample is associated with excitonic emission, while the lower energy peak is associated with recombination through impurity levels.⁹ A combination of band-acceptor and donor-acceptor-pair recombination likely accounts for the structure in the impurity-related peak. The microwaves quench the luminescence by more than one order of magnitude.

The PL and MMPL spectra of two samples with different epitaxial structures are shown in Fig. 2(a) and 2(b). The heterostructures were grown by organometallic chemical vapor deposition under conditions similar to structures reported earlier and are therefore expected to have low recombination velocities at the GaAs/GaInP interfaces.¹⁰ Diagrams of the epitaxial structures (single and double heterostructures) ac-

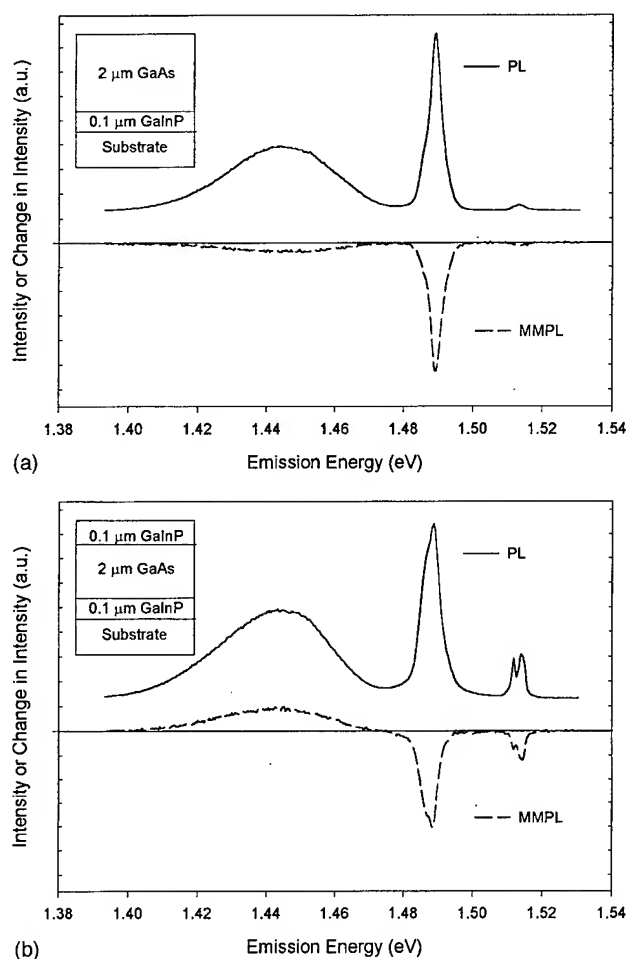


FIG. 2. PL and MMPL from the nominally identical GaAs layers in two different structures (structures also shown in figure). The scales are the same for the MMPL and PL spectra shown on each plot. The MMPL from the GaAs with the exposed surface (a) is entirely negative while that from the buried layer (b) has a significant positive component. This indicates a larger surface recombination velocity in the sample shown in (a).

company the plots. The two highest energy peaks are assigned to the same luminescence processes as those discussed in Fig. 1. The origin of the lowest energy emission, not seen for the sample depicted in Fig. 1, is not known. Two possible sources of this "extra" peak are deep impurities throughout the GaAs layer of the heterostructures and interface states associated with the GaAs/GaInP interface, which is not present in the sample whose spectra are depicted in Fig. 1. These differences are not surprising considering that the samples were grown in different laboratories by different techniques. The two samples in Fig. 2, on the other hand, were grown to be identical except for the presence of the capping layer and have very similar PL spectra. The fact that their MMPL spectra are very different will be the major comparison of this work. Note that the MMPL spectra, by our convention, are the difference between the PL spectra with the microwaves "on" and those with microwaves "off." This quantity is simply the output of a lock-in amplifier tuned to be in phase with the on-cycle of the micro-

waves. Hence, the MMPL spectrum for the situation depicted in Fig. 1 (not shown) is just the negative of the PL spectrum. The major qualitative difference between the spectra shown in Fig. 2 is that for the sample with the GaAs surface exposed, the MMPL signal is entirely negative and for the other sample the MMPL signal has a significant positive component at lower emission energy. This difference indicates that for the "bare" GaAs the microwaves enhance nonradiative recombination, while in the "capped" sample a redistribution of carriers among the various competing radiative transitions is the most significant effect. This qualitative difference has been observed for many samples.

III. DISCUSSION

Clearly, for the samples with an exposed GaAs surface [Figs. 1 and 2(a)] when the microwaves are applied the non-radiative lifetime of minority carriers becomes short compared to the radiative lifetime, and the luminescence efficiency decreases drastically. From Eq. (3), this implies that both D/d^2 and S/d become large compared to the reciprocal of the radiative lifetime. Since D/d^2 and the radiative lifetime are bulk properties which are, by design, similar for all the samples used in this study, we will postulate the existence of a relationship between the two for all samples used (with the microwaves "on") for some of the estimates made in this article:

$$\tau_r \gg \frac{d^2}{\pi^2 D}. \quad (4)$$

The MMPL and PL signals may be integrated to reveal the net effect of applying the microwave electric fields on the luminescence intensity in the spectral range shown for each of the samples depicted in Fig. 2. In the case of the double heterostructure [Fig. 2(b)], the integral of the MMPL signal is, to within experimental uncertainty, zero. This shows that the microwaves produce no measurable effect on the integrated luminescence intensity. The contrast between Figs. 2(b) and 2(a) indicates a much higher recombination velocity at the exposed GaAs surfaces than at the GaAs/GaInP interface. However, it is expected from the bare GaAs layers studied that diffusion to the GaAs/GaInP interface is significantly enhanced when the microwaves are applied. For the nonradiative lifetime not to change drastically due to the presence of the microwaves the following condition must be satisfied with both the microwaves "on" and "off:"

$$\frac{d}{2S} \gg \frac{d^2}{\pi^2 D}. \quad (5)$$

It is not clear that the thickness of the GaAs layer ($2 \mu\text{m}$) is the appropriate value for d since the entire layer may not be active, but since this distance is only a factor of two greater than the absorption length for 5145 \AA light in GaAs, this distinction will be ignored in the numerical estimates that follow. In this case Eq. (5) places an upper bound on the interface recombination velocity

$$S < \frac{\pi^2 D}{2d}, \quad (6)$$

where D refers to the (microwaves "off") 5 K minority carrier diffusion constant. The samples shown in Fig. 2 are p type, lightly doped with Zn. The acceptor concentration may be related to the relative strengths of the impurity-related and excitonic emissions.¹¹ Based on this fact and a comparison of the PL from this sample to that of a previous series with a wide range in known doping levels,³ we estimate the impurity concentration to be 10^{16} cm^{-3} . From this we can estimate an impurity-scattering-limited mobility of the minority electrons of $10^3 \text{ cm}^2/\text{V s}$, which is relatively insensitive to the number of electrons present.¹² This gives a minority carrier diffusion constant of $0.1 \text{ cm}^2/\text{s}$. Using these values, Eq. (6) constrains the interface recombination velocity to be less than 10^4 cm/s , in agreement with previous measurements on similar samples which indicated the GaInP/GaAs interface is quite high quality.¹⁰

For the "bare" GaAs layer shown in Fig. 2(a) the situation is very different. The luminescence is quenched to 72% of its original value upon exposure to the microwave electric fields. The same microwave power produced the nearly complete quenching depicted in Fig. 1. It is difficult to quantify the surface recombination velocity from this measurement without having another measurement of the quantum efficiency of the sample. However, the PL intensity as measured by our system indicates a greater efficiency (in both the absence and presence of the microwaves) than the sample shown in Fig. 1. We can therefore speculate that the surface recombination velocity is less for the sample whose spectrum is shown in Fig. 2(a) than for the one whose spectrum is shown in Fig. 1.

Any comparison between the samples depicted in Figs. 1 and 2(a) is less clear than a comparison between the two samples shown in Fig. 2 due to the differences between the PL spectra shown in Fig. 1 and those shown in Fig. 2.¹³ In order to attribute the differences in the MMPL spectra shown in Fig. 2 to differences between the bare GaAs surface and the GaAs/GaInP interface, it is important that the two samples exhibit similar bulk properties. Specifically, we are able to confirm that the (nonradiative) recombination velocity of the bare GaAs surface is larger than that of the GaAs/GaInP interface for these two samples.

IV. CONCLUSIONS

Microwave modulated photoluminescence can be used to probe both buried interfaces and exposed surfaces of any

layer of a structure that luminesces. The technique requires minimal sample preparation. Qualitative information about the nonradiative recombination at surfaces and interfaces may be obtained from a comparison of the luminescence quenching effect of the microwaves between samples with similar bulk lifetimes. Quantitative information may also be extracted if other sample parameters are known. A variable amount of microwave induced luminescence quenching from sample to sample, ranging from zero quenching to a nearly 100% effect, is indicative of the ability of MMPL to give relative information about recombination velocities in samples with different interfaces and similar bulk properties.

ACKNOWLEDGMENTS

The authors wish to thank D. S. Katzer at the Naval Research Laboratory for supplying them with additional samples. This research was supported by the National Renewable Energy Laboratory under Contract No. XAD3121142 and by the Office of Naval Research under Contract No. N000139410941.

¹I. Viohl, Ph.D. thesis, University of Utah, 1991.

²M. C. DeLong, I. Viohl, W. D. Ohlsen, P. C. Taylor, and J. M. Olson, *Phys. Rev. B* **43**, 1510 (1991).

³C. E. Inglefield, M. C. DeLong, P. C. Taylor, and W. A. Harrison, *Phys. Rev. B* (submitted).

⁴M. C. DeLong, W. D. Ohlsen, I. Viohl, X. Yin, P. C. Taylor, D. Sengupta, G. E. Stillman, J. M. Olson, and W. A. Harrison, *Phys. Rev. B* **48**, 5157 (1993).

⁵F. P. Wang, B. Monemar, and M. Ahlström, *Phys. Rev. B* **39**, 11195 (1989); H. Weman, M. Godlewski, and B. Monemar, *ibid.* **38**, 12525 (1988).

⁶*Minority Carriers in III-V Semiconductors: Physics and Applications*, edited by R. K. Ahrenkiel and M. S. Lundstrom (Academic, San Diego, 1993).

⁷G. Binning, H. Rohre, C. Gerber, and E. Wibel, *Phys. Rev. Lett.* **49**, 57 (1982); G. Binning, C. F. Quate, and C. Gerber, *ibid.* **56**, 930 (1986).

⁸M. Boulou and D. Bois, *J. Appl. Phys.* **48**, 4713 (1993).

⁹D. J. Ashen, P. J. Dean, D. T. J. Hurle, J. B. Mullin, A. M. White, and P. D. Greene, *J. Phys. Chem. Solids* **36**, 1041 (1975).

¹⁰J. M. Olson, R. K. Ahrenkiel, D. J. Dunlavy, Brian Keyes, and A. E. Kibbler, *Appl. Phys. Lett.* **55**, 1208 (1989).

¹¹Z. H. Lu, M. C. Hanna, D. M. Szymy, E. G. Oh, and A. Majerfeld, *Appl. Phys. Lett.* **56**, 177 (1990).

¹²D. C. Look, *Electrical Characterization of GaAs Materials and Devices* (Wiley, Chichester, 1989).

¹³The fact that the quenching effect shown in Fig. 2(a) is not as large as that shown in Fig. 1 may be due to a failure of the assumption given in Eq. (4). That is, the difference could arise from a difference in bulk properties rather than surface properties. This speculation is made more likely by the fact that the relative intensity of the excitonic emission is less in the samples shown in Fig. 2 than that in Fig. 1. This difference in relative intensities indicates the presence of more impurities in the samples shown in Fig. 2 than that depicted in Fig. 1, which could make D smaller, increasing the quantity on the RHS of Eq. (4).

Surface and interface effects on ellipsometric spectra of crystalline Si

K. A. Bell,^{a)} L. Mantese, U. Rossow,^{b)} and D. E. Aspnes

Physics Department, North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 28 February 1997; accepted 21 April 1997)

We present the first systematic investigation of the differences among reference-quality ellipsometrically measured pseudodielectric function $\langle\epsilon\rangle$ spectra of crystalline Si, which are nominally used to approximate the bulk dielectric function of this material. In addition to the expected influence of residual overlayers, we identify surface-local-field and energy-derivative effects, the latter representing shifts between bulk and measured critical point energies, as well as changes in excited-carrier lifetimes due to the surface. Model calculations indicate that these four effects account for nearly all differences among spectra studied, although a second-energy-derivative component appears at the E_1 transition in some cases. The isotropic contribution to the surface-local-field effect is observed for the first time. © 1997 American Vacuum Society. [S0734-211X(97)07904-3]

I. INTRODUCTION

As semiconductor devices become smaller, the need for accurate reference spectra to interpret optical data in terms of structural and material properties becomes more critical. These spectra are usually obtained from complex pseudodielectric function spectra, $\langle\epsilon\rangle = \langle\epsilon_1\rangle + i\langle\epsilon_2\rangle$, where $\langle\epsilon\rangle$ is calculated from the ellipsometrically measured complex reflectance ratio in the two-phase (substrate/ambient) model without regard for surface overlayers. These reference spectra may be $\langle\epsilon\rangle$ spectra of samples for which considerable care has been expended to eliminate and/or prevent the formation of overlayers, or they may be calculated from $\langle\epsilon\rangle$ spectra by compensating for the effect of overlayers usually in the three-phase (substrate/overlayer/ambient) model. In recent publications, several reference-quality spectra for crystalline (*c*-) Si were reported and compared^{1,2} to previously published reference-quality spectra.³ Also, differences in optically modeled film thicknesses resulting from the use of different reference spectra have been noted.⁴ The mechanisms that give rise to these differences, however, have not been addressed. Here, we report the first systematic investigation of differences among nominally identical $\langle\epsilon\rangle$ spectra for *c*-Si. We find unexpectedly large differences of up to 4.5% at the (E'_0, E_1) complex (3.32 and 3.4 eV, respectively) and 3.3% at the E_2 transition (4.2 eV) in these nominally *bulk* dielectric functions. Differences include shifts in apparent critical point energies even for seemingly identical H-terminated surfaces.

We find that these differences can be grouped into four categories: (1) expected residual-overlayer effects due to the presence of oxides or microscopic roughness on the surface; (2) the surface-local-field (SLF) effect;⁵⁻⁷ and energy-derivative effects⁶⁻⁸ that represent (3) shifts ΔE_g of critical point energies E_g , and (4) shifts $\Delta\Gamma$ of broadening parameters Γ from one spectrum to the next (and, therefore, prob-

ably from their true bulk values). In particular, we find a large *isotropic* SLF difference between surfaces terminated with SiO₂ and H. Isotropic SLF contributions have not been observed previously. The quantities ΔE_g and $\Delta\Gamma$ depend on both sample termination and orientation. Differences in the broadening parameter can be understood in terms of increased scattering at surfaces that cannot be prepared to the same level of quality as Si(111)-H. The reasons for differences among measured critical point energies with chemical termination are less obvious, but presumably result from a combination of a difference in surface potentials and the localization of optically excited carriers, as discussed more completely by Mantese *et al.*⁹

The SLF and derivative contributions indicate that systematic differences among $\langle\epsilon\rangle$ spectra cannot be corrected by three-phase model calculations alone, but that more fundamental mechanisms are present. These results further show that accurate modeling requires the use of a reference spectrum obtained on a sample with the appropriate chemical termination. Also, differences in spectra obtained for nominally identical Si(111)-H surfaces show that care must be taken to eliminate steps and other imperfections in samples intended for reference-quality spectra. Finally, the mechanisms discussed must be taken into account in theoretical calculations of surface-optical properties if the full diagnostic power of surface-optical techniques is to be realized.

II. EXPERIMENT AND DATA

Samples measured here were nominally on-axis ($\pm 0.5^\circ$), *n*-type Si(111) and Si(110) wafers grown and chemomechanically polished by Western Electric Corporation. Resistivities were $\sim 800 \Omega \text{ cm}$ for Si(111) and $24 \Omega \text{ cm}$ for Si(110). The Si(111) samples were prepared by a three-step RCA clean followed by a 5 min immersion in NH₄F (40%) to produce a stable, atomically smooth, H-terminated surface.¹⁰ Our specific procedure is described by Yasuda and Aspnes.¹¹ Samples were never allowed to dry in air at any time between treatments. Following the NH₄F dip, samples were blown dry with nitrogen. Oxidized surfaces were pre-

^{a)}Electronic mail: bell@unity.ncsu.edu

^{b)}Present address: Institut für Physik, Technische Universität, Ilmenau, Germany D-98684.

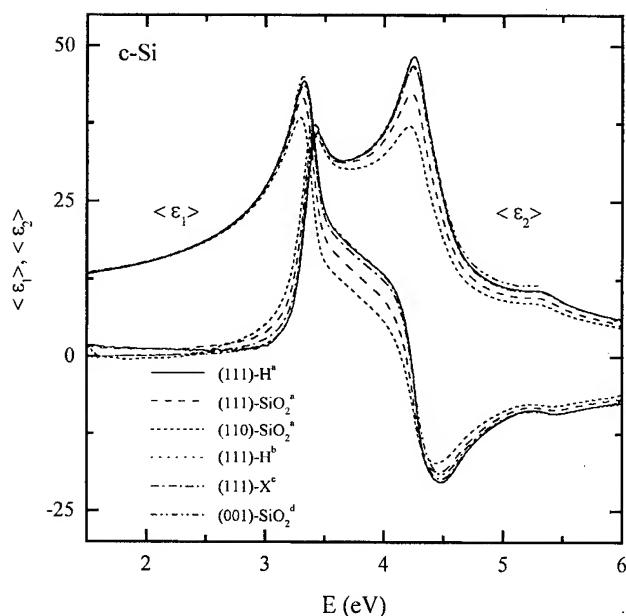


FIG. 1. Representative pseudodielectric function spectra of *c*-Si for various orientations and chemical terminations. The subscript "a" indicates data obtained here; subscripts "b, c, and d" indicate data obtained from Refs. 11, 3, and 1, respectively. Here, and in subsequent figures, the notations -H, -SiO₂ and -X refer to the chemical termination of the respective surface.

pared by immersing the Si(111)-H surfaces for 30 s in H₂O₂. The Si(110) sample was air-oxidized and used as is, except for a methanol (MeOH) rinse prior to measurement. Samples were mounted vertically on a vacuum chuck in a windowless glass cell that allowed the surface to be accessed optically while maintained in a dry N₂ ambient to ensure chemical stability. Anisotropy data were obtained by measuring the Si(110)-SiO₂ sample at 45° ± 5° increments relative to the plane of incidence, where the initial measurement was obtained along a principal axis of surface as determined by reflectance-difference spectroscopy (RDS).

Our data were obtained from 1.5 to 6.0 eV using a rotating analyzer ellipsometer (RAE); instrumentation¹² and experimental techniques³ are described in detail elsewhere. Data acquisition required 9 min per spectrum and typically began 30 min after sample preparation. The Si(111) samples were measured for both H and SiO₂ terminations prepared as described above. Consistency was monitored by the height of the E₂ peak in $\langle \epsilon_2 \rangle$ at 4.2 eV. The maximum peak height for our Si(111)-H surfaces was 48.2 ± 0.1; the highest value reported to date is 48.3 ± 0.1, obtained by Yasuda and Aspnes.¹¹ Once oxidized, the initial high-quality Si(111)-H surfaces could not be recovered by treatment with either (1:30) dilute HF:H₂O or (40%) NH₄F. In particular, the NH₄F etch tended to roughen and degrade air-exposed surfaces, in agreement with previous scanning tunneling microscopy¹³ and atomic force microscopy¹⁴ results.

Representative spectroscopic ellipsometry (SE) data for our Si(111)-H, Si(111)-SiO₂, and Si(110)-SiO₂ surfaces are shown in Fig. 1. Also shown are previously reported reference-quality spectra obtained from the literature:

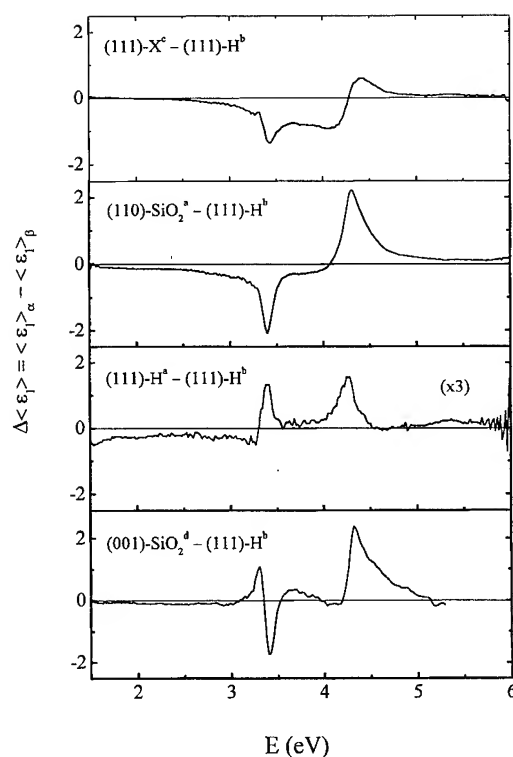


FIG. 2. Representative differences $\Delta\langle\epsilon_1\rangle$ between selected pairs of spectra in Fig. 1 as indicated.

Si(111)-H prepared as described above;¹¹ Si(111)-X prepared by a bromine-methanol followed by a subsequent chemical treatment,³ where X denotes the fact that the surface termination was unknown; and Si(001)-SiO₂ prepared by a HF etch followed by 1 h oxidation in air, where the effect of the native oxide was removed mathematically using the three-phase model.¹ Though these reference spectra are nominally identical, we find substantial systematic differences that are best illustrated by subtracting one spectrum from another. Results for the real parts of these differences, $\Delta\langle\epsilon_1\rangle$, are shown in Fig. 2 for selected pairs. These differences and the mechanisms giving rise to them are discussed in detail below.

III. THEORY

In the following, we develop mathematical models to describe the $\Delta\langle\epsilon_1\rangle$ line shapes of Fig. 2 and show how they can be used to interpret observed systematic differences in SE data. Instrumentation artifacts can be eliminated as contributing factors since the line shapes resulting, for example, from errors in calibration constants or the angle of incidence, ϕ , do not resemble the data shown.

As mentioned above, the pseudodielectric function $\langle\epsilon\rangle$ is the two-phase model representation of the ellipsometrically measured complex reflectance ratio, $\rho = r_p/r_s$, where r_p and r_s are the complex reflectances of *p*- and *s*-polarized light, respectively. Specifically

$$\frac{\langle \epsilon \rangle}{\epsilon_a} = \sin^2 \phi + \sin^2 \phi \tan^2 \phi \left(\frac{1-\rho}{1+\rho} \right)^2, \quad (1)$$

where ϵ_a is the ambient dielectric function usually assumed to be 1. In the absence of overlayers $\langle \epsilon \rangle = \epsilon_s$, where ϵ_s is the bulk dielectric function of the substrate.

If the sample is covered by an intentionally deposited, unintentionally accumulated, or surface-modified overlayer of dielectric function ϵ_o and thickness d , then $\langle \epsilon \rangle$ is related to ϵ_s to first order in d/λ by¹⁵

$$\langle \epsilon \rangle = \epsilon_s + \frac{4\pi i d}{\lambda} \frac{\epsilon_s(\epsilon_s - \epsilon_o)(\epsilon_o - \epsilon_a)(\epsilon_s - \epsilon_a \sin^2 \phi)^{1/2}}{\epsilon_o(\epsilon_s - \epsilon_a)}, \quad (2)$$

where λ is the wavelength of light. The difference between $\langle \epsilon \rangle$ and the true bulk dielectric function of the substrate (which may differ from ϵ_s) is the correction term $\Delta \langle \epsilon \rangle$. We consider four models for $\Delta \langle \epsilon \rangle$ which are described below.

A. Overlayer model

If the quality of preparation differs for two samples, we can expect differences in the amount of microscopic roughness (areal density of steps) or thickness of residual overlayers on their surfaces, such that $\Delta \langle \epsilon \rangle_{\text{overlayer}}$ originates from the second term on the right-hand side of Eq. (2). If $|\epsilon_s| \gg |\epsilon_o| \approx \epsilon_a$, this term simplifies to

$$\Delta \langle \epsilon \rangle_{\text{overlayer}} = (\langle \epsilon \rangle - \epsilon_s)_{\text{overlayer}} \approx \frac{4\pi i d}{\lambda} \epsilon_s^{3/2} \left(1 - \frac{\epsilon_a}{\epsilon_o} \right). \quad (3)$$

For $|\epsilon_o| \gg \epsilon_a$, the term in parentheses drops out and Eq. (3) is independent of the overlayer dielectric function. We considered both fused SiO₂, taken from Malitson,¹⁶ and microscopic roughness, approximated as a Bruggeman mix of 50% c-Si and 50% SiO₂, as representative dielectric functions for ϵ_o . In the former case, ϵ_o is real such that the scaling term only affects the value of d and can be neglected. Though roughness does a slightly better job fitting the data than SiO₂ in some cases, as is discussed below, we find no clear distinction between results obtained using either to approximate ϵ_o . For consistency, however, each difference spectrum was fit using the roughness overlayer line shape shown in Fig. 3(a) for ϵ_s equal to the dielectric function of c-Si from Ref. 11, $\epsilon_a = 1$ and $d = 1 \text{ \AA}$.

B. Surface-local-field (SLF) model

The SLF effect originates from a modification of bulk screening due to the surface, and is known to give a large contribution to certain anisotropy spectra.⁵⁻⁷ In the contact exciton model,¹⁷ bulk screening is described according to $\epsilon_s = \epsilon_{so} / (1 + \tilde{g} E^2 \epsilon_{so})$, where ϵ_{so} is the unscreened bulk dielectric function, and $|\epsilon_{so}|, |\epsilon_s| \gg 1$. Also, $\tilde{g} < 0$ is the contact exciton interaction or "screening" parameter that, in general, is expected to be tensorial with different values normal and parallel to the surface and, for vicinal (001) and (111) surfaces, even within the surface plane. Over a narrow energy range, we can use the SLF screening parameter $g = \tilde{g} E^2$ and treat g as a constant.⁶ Supposing that $g \rightarrow g$

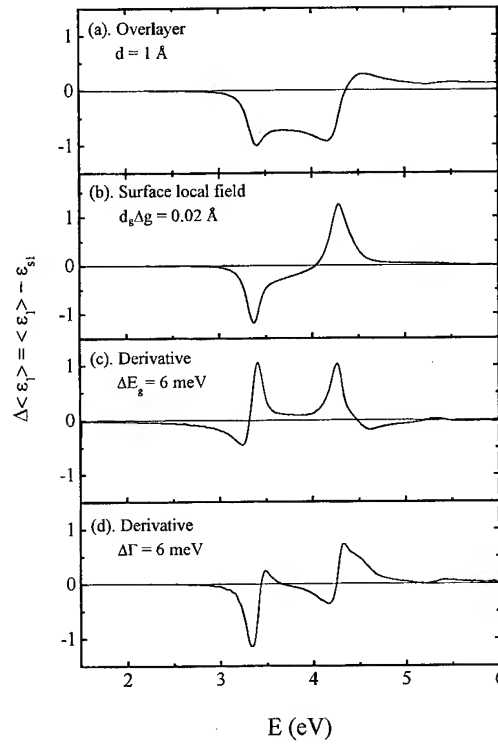


FIG. 3. Model line shapes: (a) microscopic roughness overlayer model for $d = 1 \text{ \AA}$; (b) surface-local-field model for $d_g \Delta g = 0.02 \text{ \AA}$; (c) derivative with respect to critical point energy for $\Delta E_g = 6 \text{ meV}$; and (d) derivative with respect to broadening parameter for $\Delta \Gamma = 6 \text{ meV}$.

+ Δg over a depth $d_g \ll \lambda$ near the surface, we have $\epsilon_o \approx \epsilon_s + \Delta g \partial \epsilon_s / \partial g$, in which case, $\Delta \langle \epsilon \rangle_{\text{SLF}}$ again comes from the second term on the right-hand side of Eq. (2) for $d = d_g$,

$$\Delta \langle \epsilon \rangle_{\text{SLF}} = (\langle \epsilon \rangle - \epsilon_s)_{\text{SLF}} \approx \frac{4\pi i d_g}{\lambda} \Delta g \epsilon_s^{5/2}. \quad (4)$$

The line shape for $\Delta \langle \epsilon \rangle_{\text{SLF}}$ is shown in Fig. 3(b) for c-Si for $d_g \Delta g = 0.02 \text{ \AA}$.

C. Derivative models

In the absence of directional perturbations such as uniaxial strain, derivativelike features in optical spectra arise from shifts ΔE_g and/or $\Delta \Gamma$ of apparent critical point energies E_g and broadening parameters Γ from their bulk values. We interpret these shifts as bulk effects driven by a localization of the optically excited carriers near the surface.⁹ Also, since Γ is inversely proportional to the lifetime of the excitation process, measurements of $\Delta \Gamma$ between spectra indicate differences in the excited carrier lifetimes; rougher surfaces are expected to have shorter lifetimes than smoother surfaces due to scattering.

The associated difference line shapes, $\Delta \langle \epsilon \rangle_{\Delta E_g}$ and $\Delta \langle \epsilon \rangle_{\Delta \Gamma}$, follow by noting that near a critical point, ϵ_s can be written as

$$\epsilon_s(E) \approx B(E) + L(E - E_g + i\Gamma), \quad (5)$$

where $B(E)$ and $L(E - E_g + i\Gamma)$ represent the contributions of the slowly varying background and the critical point, respectively. In these cases, $\Delta\langle\epsilon\rangle$ arises from the first term on the right-hand side of Eq. (2) for $d = 0$. To first order in ΔE_g and $\Delta\Gamma$ we have

$$\Delta\langle\epsilon\rangle_{\Delta E_g} = (\langle\epsilon\rangle - \epsilon_s)_{\Delta E_g} = -\Delta E_g \frac{d\epsilon_s}{dE}, \quad (6)$$

$$\Delta\langle\epsilon\rangle_{\Delta\Gamma} = (\langle\epsilon\rangle - \epsilon_s)_{\Delta\Gamma} = i\Delta\Gamma \frac{d\epsilon_s}{dE}. \quad (7)$$

The line shapes corresponding to $\Delta\langle\epsilon_1\rangle_{\Delta E_g}$ and $\Delta\langle\epsilon_1\rangle_{\Delta\Gamma}$ are shown in Figs. 3(c) and 3(d), respectively, for $\Delta E_g = \Delta\Gamma = 6$ meV.

Since the differences among reference spectra are small (less than 10% of $\langle\epsilon\rangle$), the correction terms in Eqs. (3), (4), (6), and (7) can be used additively to model differences among SE data.

IV. RESULTS AND DISCUSSION

Using least-squares regression, we determined the relative contributions of each of the four models described above to difference spectra, such as those shown in Fig. 2. The real parts of these differences, $\Delta\langle\epsilon_1\rangle$, are shown as the solid lines in Fig. 4. The dashed curves represent the best fit to each line shape. We note that model fits to the real and imaginary parts of these differences were consistent to within three mean-square deviations above 3.4 eV with few exceptions. Below this energy, our $\langle\epsilon_2\rangle$ data are less accurate due to the poor sensitivity of the RAE to small values of $\langle\epsilon_2\rangle$.

Since SLF and derivative model contributions to the (E'_0, E_1) complex and E_2 transition were different, fits were done separately from 1.5 to 4.0 eV for (E'_0, E_1) and 3.6 to 6.0 eV for E_2 . Differences in SLF contributions arise from changes in the screening parameter g , which may vary according to the energy range of interest; differences in ΔE_g and $\Delta\Gamma$ derivative models arise since each critical point and/or broadening parameter may experience different shifts. This is also true within the complex. For example, it has been shown that E'_0 and E_1 for *c*-Si shift by 40 and 70 meV, respectively, as the temperature is raised from ~ 0 to 283 K.¹⁸ These derivative effects make it difficult to fit the line shape in the (E'_0, E_1) region with our models in some cases. For example, difference spectra obtained for SiO₂- and H-terminated surfaces exhibit a second-derivative line shape, possibly due to the near-degeneracy of the E'_0 and E_1 critical points.⁹

Due to the complexity of the (E'_0, E_1) structure, we focus our attention on the E_2 region. However, the best-fit coefficients d , $d_g\Delta g$, ΔE_g , and $\Delta\Gamma$ and their uncertainties are listed in Table I for both energy ranges along with respective mean-square deviations, δ . In each case, all four coefficients were determined simultaneously to represent model correlations accurately. From 1.5 to 4.0 eV, the strongest correlations, i.e., above 70%, are between the overlayer and SLF models (80%), and between the SLF and $\Delta\Gamma$ derivative models (80%). From 3.6 to 6.0 eV, the energy range discussed

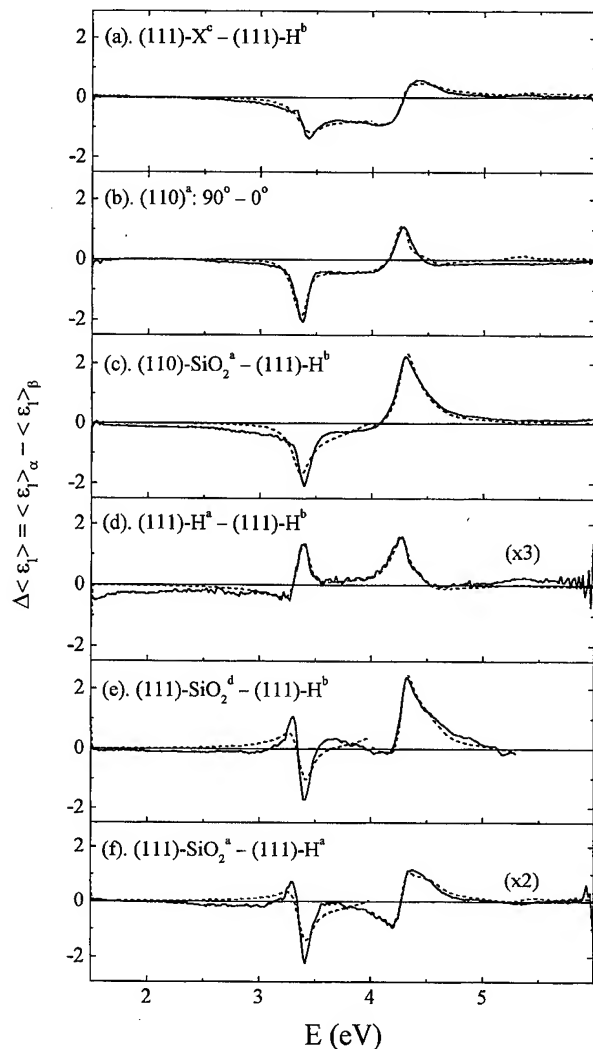


FIG. 4. Least-squares fits (dashed lines) to differences (solid lines) of the real part of the dielectric function $\Delta\langle\epsilon_1\rangle$ between selected pairs of spectra as indicated. Since model contributions to the (E'_0, E_1) complex and the E_2 transition are in principle different, fits were done over separate energy ranges: 1.5–4.0 eV for (E'_0, E_1) , and 3.6–6.0 eV for E_2 .

here, the strongest correlations are between the SLF and ΔE_g derivative models (90%), as well as between the overlayer and ΔE_g derivative models (80%).

Since no *a priori* estimates exist for the orders of magnitude of the SLF and derivative effects, these were established experimentally. Thus, the scale of $d_g\Delta g$ is set by the value 0.042 ± 0.004 Å determined for the difference between our isotropic Si(110)–SiO₂ data and that reported by Yasuda and Aspnes¹¹ for Si(111)–H; the scale of ΔE_g is set by the value 9.6 ± 1.1 meV determined for the Si(110): 90°–0° anisotropy spectrum; and the scale of $\Delta\Gamma$ is set by the value 14.1 ± 0.5 meV determined for the difference between the Jellison¹ and Yasuda and Aspnes¹¹ data. We consider first those differences that best illustrate each of the four models, then discuss implications regarding such differences for reference spectra used for material and structural modeling.

TABLE I. Least-squares results for differences between pairs of spectra shown in Fig. 1. The notation -H, -SiO₂, and -X refers to the chemical termination of the respective surface. In entries 4(c) and 4(f), the effects of the roughness overlayer were removed mathematically.

Figure	Difference spectrum	Energy (eV)	Overlayer $d(\text{\AA})$	SLF $d_g\Delta g(\text{\AA})$	Derivative E_g $\Delta E_g(\text{meV})$	Derivative Γ $\Delta\Gamma(\text{meV})$	δ
4(a)	(111)-X ^c -(111)-H ^b	1.5-4.0	1.05 ± 0.04	0.010 ± 0.002	1.6 ± 0.3	-2.6 ± 0.4	0.08
		3.6-6.0	1.00 ± 0.04	0.004 ± 0.001	1.5 ± 0.7	1.8 ± 0.3	0.08
4(b)	(110)-SiO ₂ ^a : 90°-0°	1.5-4.0	0.27 ± 0.04	0.018 ± 0.002	-0.7 ± 0.3	2.9 ± 0.4	0.08
		3.6-6.0	0.75 ± 0.07	0.000 ± 0.002	9.6 ± 1.1	-1.2 ± 0.6	0.13
4(c)	(110)-SiO ₂ ^a -(111)-H ^b (8.9 Å roughness removed)	1.5-4.0	-0.11 ± 0.09	0.042 ± 0.004	1.7 ± 0.7	-2.8 ± 0.9	0.18
		3.6-6.0	-0.29 ± 0.04	0.037 ± 0.002	-4.5 ± 0.7	5.0 ± 0.4	0.09
4(d)	(111)-H ^a -(111)-H ^b	1.5-4.0	-0.03 ± 0.03	0.002 ± 0.001	2.5 ± 0.2	-1.0 ± 0.3	0.07
		3.6-6.0	-0.07 ± 0.04	0.004 ± 0.001	1.3 ± 0.7	-0.1 ± 0.3	0.08
4(e)	(001)-SiO ₂ ^d -(111)-H ^b	1.5-4.0	-0.51 ± 0.12	0.019 ± 0.006	-4.7 ± 0.9	-4.0 ± 1.2	0.25
		3.6-5.3	-0.92 ± 0.06	0.023 ± 0.002	-8.0 ± 1.1	14.1 ± 0.5	0.15
4(f)	(111)-SiO ₂ ^a -(111)-H ^a (4.8 Å roughness removed)	1.5-4.0	0.00 ± 0.06	0.011 ± 0.003	-1.7 ± 0.4	-2.6 ± 0.6	0.12
		3.6-6.0	-0.03 ± 0.04	0.001 ± 0.002	-2.1 ± 0.7	4.9 ± 0.4	0.08

^aData obtained here.

^bReference 11.

^cReference 3.

^dReference 1.

A. Overlayer effects

The overlayer line shape appears most clearly in the difference between the Aspnes and Studna³ and Yasuda and Aspnes¹¹ data shown in Fig. 4(a). The best-fit coefficients listed in Table I for the E_2 transition give an overlayer thickness of $d = 1.00 \pm 0.04$ Å with comparatively small contributions from the other models. The sign of d indicates that the Aspnes and Studna data³ were obtained on a lower quality surface than that of Yasuda and Aspnes,¹¹ which is not surprising since the NH₄F treatment was not used in the former case. This is further supported by the value of $\Delta\Gamma = 1.8 \pm 0.3$ meV, which indicates larger scattering for the Aspnes and Studna³ sample. The relatively small value of $\Delta E_g = 1.5 \pm 0.7$ meV, however, indicates that this surface was terminated largely by H, as will be discussed below. If the overlayer is assumed to be SiO₂, we obtain $d \approx 2$ Å as was speculated by Yasuda and Aspnes.¹¹ Other than the factor of 2 in d that arises from the scaling factor in Eq. (3), as described above, there is no distinction between approximating the overlayer as microscopic roughness or SiO₂.

B. Surface-local-field effects

The SLF effect is well known in anisotropy spectroscopy, where it dominates the (normal-incidence) RD spectra of (110) Si and Ge surfaces.^{6,7} It is also present in the RD spectra of vicinal Si(001) and Si(111) surfaces off-cut toward [110], presumably due to [110]-type steps on the surface,^{7,8} and may be present in many clean semiconductor surfaces such as (2×4) GaAs (001).¹⁹ The ellipsometrically determined anisotropy spectrum for the Si(110)-SiO₂ surface, shown in Fig. 4(b), was calculated by taking the difference between spectra obtained at azimuth angles 90° and 0°, which correspond to the alignment of the plane of incidence along the respective principal axes of the surface, α and β ,

and yields the maximum anisotropy signal. This line shape strongly resembles the SLF model in Fig. 3(b), though its appearance is qualitatively different from that observed with RDS. We note that a full spectrum fit from 1.5 to 6.0 eV (not shown) gave a value of $d_g\Delta g = 0.019 \pm 0.002$ Å, where $\Delta g = g_\alpha - g_\beta$ represents screening anisotropy in the plane of the surface. However, we find a much smaller value of $d_g\Delta g = 0.000 \pm 0.002$ Å for the E_2 region alone, as shown in Table I. This dramatic decrease in $d_g\Delta g$ may be illusory since it is accompanied by a similar large change in ΔE_g from 0.1 ± 0.6 to 9.6 ± 1.1 meV. As mentioned above, these model line shapes are highly correlated over this energy range, and some compensation is undoubtedly occurring. We also note that the value of $\Delta\Gamma = -1.2 \pm 0.6$ meV indicates a difference in scattering between the 90° and 0° sample orientations due to the (110)-type steps on the surface.

More interesting is the difference between the isotropic Si(110)-SiO₂ data and the Yasuda and Aspnes¹¹ data shown in Fig. 4(c). Here, the isotropic contribution to Si(110)-SiO₂ was isolated by averaging data taken at the eight different azimuths mentioned above. Though we expected the Si(110)-SiO₂ sample to be oxide terminated, we fit the data more accurately by assuming a microscopic roughness overlayer. By mathematically removing the 8.9 Å of roughness determined by a least-squares fit to the initial difference spectrum (not shown), we highlight the residual line shape. The coefficients of the least-squares fit to this line shape are listed in Table I and show a SLF contribution of $d_g\Delta g = 0.037 \pm 0.002$ Å for the E_2 transition. In this case, Δg represents the difference between the *normal* components of the screening tensors for the Si(110)-SiO₂ and Si(111)-H surfaces. This is the first time that such differences have been observed. We also find relatively large derivative contributions in this energy range of $\Delta E_g = -4.5$

± 0.7 meV and $\Delta\Gamma = 5.0 \pm 0.4$ meV. The former can be attributed to the difference in chemical termination between these surfaces (see below). The sign of $\Delta\Gamma$ shows that Γ is larger for the Si(110)-SiO₂ surface than for Si(111)-H. This is consistent with expectations, since the Si(111)-H surface is atomically smooth and the lifetimes of excited carriers are expected to be shorter in the rougher Si(110)-SiO₂ sample due to scattering.

C. Critical point energy (ΔE_g) effects

The line shape in Fig. 3(c) is most clearly seen in the difference between the Si(111)-H spectra obtained here and that reported by Yasuda and Aspnes¹¹ shown in Fig. 4(d). As these surfaces were prepared by the same procedure, these spectra are nominally identical. The difference spectrum is small, with the dominant contribution from $\Delta E_g = 2.5 \pm 0.2$ meV for the (E'_0, E_1) complex and 1.3 ± 0.7 meV for the E_2 transition. We note that temperature effects, known to give rise to energy shifts on the order of 0.2 meV/K,²⁰ are negligible. By monitoring a test sample *in situ*, we found that changes in temperature of up to 5 K could occur due to evaporative cooling from a MeOH rinse, however, these changes vanished in less than 1 min. We further rule out bulk strain as a mechanism for the shift by using a typical value of 10 eV for the hydrostatic volume deformation potential²¹ in strain calculations. A shift of 2.5 meV would require a 0.025% strain, which is not likely to be caused by variations in chemical preparations between samples. Instead, these shifts indicate that the E'_0 , E_1 , and E_2 bulk *c*-Si critical point energies determined from optical spectra are apparent only and not equal to their true bulk values, in agreement with previous conclusions obtained by RDS.^{7,8} This is discussed more fully in Ref. 9.

Larger ΔE_g contributions are found in the differences between H- and SiO₂-terminated surfaces. For example, the difference between the Jellison¹ and Yasuda and Aspnes¹¹ data, shown in Fig. 4(e), exhibits sharp derivative-type features near bulk critical point energies. As mentioned above, we observe a strong second-derivative contribution to the line shape for the (E'_0, E_1) complex, which suggests that our set of models is incomplete or that the (E'_0, E_1) complex cannot be treated as a single critical point. However, the obvious derivative nature of these spectra and the relatively large value of $\Delta E_g = -8.0 \pm 1.1$ meV for the E_2 transition indicates that some, and possibly all, critical point energies are shifted from their true bulk values. We also find an isotropic SLF effect between these surfaces with a value of $d_g \Delta g = 0.023 \pm 0.002$ Å. The negative value of d in Table I reflects some discrepancies between our overlayer model and the model Jellison used to correct his data.¹

The Fig. 4 results are the first indication that chemical termination plays a role in reference SE data, even after compensating for overlayer effects. To investigate this further, we examine differences between spectra obtained for H- and SiO₂-terminated Si(111) samples, where the SiO₂ layer was formed by dipping the Si(111)-H surfaces in H₂O₂ as described above. As expected, a least-squares fit to the initial

difference spectrum (not shown) indicated a dominant overlayer contribution due to oxide on the surface. The data were better fit using the roughness overlayer model, which indicates that the H₂O₂ treatment roughened the surface. We isolate the more fundamental mechanisms involved by mathematically removing 4.8 Å of microscopic roughness. The remaining line shape is shown in Fig. 4(f). Again, we find a second derivative contribution to the (E'_0, E_1) complex but a better fit to the E_2 transition with $\Delta E_g = -2.1 \pm 0.7$ meV.

D. Lifetime ($\Delta\Gamma$) effects

From the difference spectra discussed above, we have already seen that differences in surface quality give rise to differences in Γ between samples. This was seen strongly in the difference between Si(110)-SiO₂ and Si(111)-H shown in Fig. 4(c), and is also found in the line shapes shown in Figs. 4(e) and 4(f). In Fig. 4(e), the best fit to the line shape gives a value of $\Delta\Gamma = 14.1 \pm 0.5$ meV for the E_2 transition, which indicates a substantially larger Γ for Si(001)-SiO₂ than for Si(111)-H. This is consistent with the fact that Si(001) surfaces cannot be chemically prepared to the same degree of perfection as Si(111)-H surfaces and should, therefore, exhibit larger amounts of scattering. In Fig. 4(f), the difference between our Si(111)-SiO₂ and Si(111)-H spectra, the value of $\Delta\Gamma$ measured is 4.9 ± 0.4 meV, indicating that $\Delta\Gamma$ is a measure of the difference in surface quality between SiO₂- and H-terminated Si(111). It seems reasonable that the data giving rise to relatively large values of $\Delta\Gamma$ are better fit using microscopic roughness as ϵ_o in the overlayer model rather than SiO₂. Though this is the case for the data shown in Figs. 4(c) and 4(f), we find no clear distinction for the difference spectrum shown in Fig. 4(e).

V. SUMMARY

This article serves as the first systematic investigation of differences among reference-quality ellipsometric data for *c*-Si. The above results are general and are found not only between H- and SiO₂-terminations on the same sample, but also on different samples. This observation is the basis for our above statement that the lack of a ΔE_g derivative component implies that the Aspnes and Studna³ surface was nearly H-terminated. We conclude, therefore, that $\Delta\Gamma$ is a general measure of surface quality while ΔE_g gives information about chemical termination, possibly influenced by orientation. Finally, the isotropic SLF effect that appears systematically among these difference spectra indicates that the normal component of the screening tensor also depends on sample orientation and termination.

We recognize that more analysis needs to be done to clearly distinguish between microscopic roughness and SiO₂ overlayer contributions to the differences discussed, as well as to develop more representative models for the (E'_0, E_1) complex that include second-energy-derivative effects. This work is currently in progress. However, from the results presented here, it is clear that accurate modeling of SE data requires a reference spectrum that is consistent with

respect to orientation and chemical termination as far as possible. To this end, we note that energy shifts recently reported in SE data^{2,22} may indicate inconsistencies between the samples measured and the reference data used for optical modeling. We suggest that a reference spectrum more nearly representative of the starting surfaces may give different results and that the shifts observed may be due to localization, as discussed in Ref. 9.

ACKNOWLEDGMENTS

The authors would like to acknowledge support of this work by ONR and DARPA. U. Rossow would like to thank the Alexander von Humbolt Foundation for further support.

¹G. E. Jellison, Jr., Opt. Mater. **1**, 41 (1992).

²V. Nayar, W. Y. Leong, C. Pickering, A. J. Pidduck, R. T. Carline, and D. J. Robbins, Appl. Phys. Lett. **61**, 1304 (1992).

³D. E. Aspnes and A. A. Studna, Phys. Rev. B **27**, 985 (1983).

⁴W. A. McGahan and J. A. Woollam, in *Proceedings of the International Workshop on Semiconductor Characterization: Present Status and Future Needs*, edited by W. M. Bullis, D. G. Seiler, and A. Diebold (AIP, Woodbury, 1995), p. 433.

⁵W. L. Mochán and R. G. Barrera, J. Phys. (Paris) Colloq. Suppl. **45**, C5-207 (1984).

⁶D. E. Aspnes and A. A. Studna, Phys. Rev. Lett. **54**, 1956 (1985); D. E. Aspnes, J. Vac. Sci. Technol. B **3**, 1498 (1985).

⁷T. Yasuda, D. E. Aspnes, D. R. Lee, C. H. Bjorkman, and G. Lucovsky, J. Vac. Sci. Technol. A **12**, 1152 (1994).

⁸T. Yasuda, L. Mantese, U. Rossow, and D. E. Aspnes, Phys. Rev. Lett. **74**, 3431 (1995); L. Mantese, U. Rossow, and D. E. Aspnes, Appl. Surf. Sci. **107**, 35 (1996).

⁹L. Mantese, K. A. Bell, U. Rossow, and D. E. Aspnes, J. Vac. Sci. Technol. B **15**, 1196 (1997).

¹⁰G. S. Higashi, R. S. Becker, Y. J. Chabal, and A. J. Becker, Appl. Phys. Lett. **58**, 1656 (1991).

¹¹T. Yasuda and D. E. Aspnes, Appl. Opt. **33**, 7435 (1994).

¹²D. E. Aspnes and A. A. Studna, Appl. Opt. **14**, 220 (1975); Rev. Sci. Instrum. **49**, 291 (1978).

¹³G. S. Hsiao, J. A. Virtanen, and R. M. Penner, Appl. Phys. Lett. **63**, 1119 (1993).

¹⁴T. Suzuki and S. Adachi, Jpn. J. Appl. Phys. **1** **33**, 5599 (1994).

¹⁵D. E. Aspnes, in *Optical Properties of Solids: New Developments*, edited by B. O. Seraphin (North-Holland, Amsterdam, 1976), p. 799.

¹⁶I. H. Malitson, J. Opt. Soc. Am. **55**, 1205 (1965).

¹⁷Y. Toyozawa, M. Inoue, T. Inui, M. Okazaki, and E. Hanamura, J. Phys. Soc. Jpn. **22**, 1337 (1967).

¹⁸A. Daunois and D. E. Aspnes, Phys. Rev. B **18**, 1824 (1978).

¹⁹U. Rossow, L. Mantese, and D. E. Aspnes, in *Proceedings of the 23rd International Conference on the Physics of Semiconductors*, edited by M. Scheffler and R. Zimmermann (World Scientific, Singapore, 1996), p. 831.

²⁰P. B. Allen and M. Cardona, Phys. Rev. B **27**, 4760 (1983).

²¹P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 1996), p. 118.

²²N. V. Nguyen, D. Chandler-Horowitz, P. M. Amirtharaj, and J. G. Pellegrino, Appl. Phys. Lett. **64**, 2688 (1994).

Reflectance-difference studies of interface-formation and initial-growth processes in ZnSe/GaAs(001) heteroepitaxy

T. Yasuda,^{a),b)} K. Kimura,^{c)} S. Miwa,^{c)} L. H. Kuo,^{c)} A. Ohtake,^{c)} C. G. Jin,^{d)} K. Tanaka,^{b),d)} and T. Yao^{b),e)}

Joint Research Center for Atom Technology (JRCAT), 1-1-4 Higashi, Tsukuba 305, Japan

(Received 29 January 1997; accepted 16 April 1997)

In situ reflectance-difference studies of initial stages of ZnSe growth on GaAs(001) surfaces are reported. ZnSe layers with thicknesses less than 4 nm were grown by molecular beam epitaxy in both layer-by-layer and island-growth modes. It is found that the reflectance-difference spectra for this thickness range are clearly dependent on ZnSe thickness as well as on the growth mode. This observation indicates that the surface electronic structure develops with thickness and in a manner sensitive to the mid- or long-range order of the surface. The interface-induced peak at 2.8 eV, intensity of which is correlated with the extent of the interfacial Ga-Se bond formation, is pinned during the course of growth, which indicates that atomic rearrangement or mixing at the interface is minimal once several monolayers of ZnSe are deposited on GaAs. © 1997 American Vacuum Society. [S0734-211X(97)07704-4]

I. INTRODUCTION

Optical diagnostics of crystal growth has been attracting increasing attention, not only because it can be performed in nonvacuum environments, but also because it provides us with information about the electronic state which is complementary to the structural information obtained by commonly used diffraction techniques. Reflectance-difference spectroscopy (RDS), which measures polarization dependence, or anisotropy, in normal-incidence reflectance, is among the surface-sensitive optical techniques suitable for *in situ* characterization for materials processing.^{1,2} Since its first demonstration by Aspnes and Studna in 1985,³ there have been many RDS studies of surface-induced anisotropy. It was only recently, however, that the same technique was applied to heterostructures where optical anisotropies of various origins are possible.⁴⁻¹¹ For the ZnSe/GaAs system, the surface-induced anisotropy was the primary concern of recent studies by several groups. Anisotropies of the other origins were not examined in detail in their reports.¹²⁻¹⁴

In our previous publications, we demonstrated that *in situ* RDS can simultaneously measure the surface- and interface-induced optical anisotropies of a heterostructure if the acquired data are properly analyzed.^{4,9} An application of this technique to ZnSe/GaAs(001) heterostructures, where the ZnSe epilayer thickness, L_{ZnSe} , was varied from 3 to 40 nm, showed that both surface- and interface-induced anisotropies are essentially independent of L_{ZnSe} . The RD responses for the thin-film regime ($L_{\text{ZnSe}} < \text{a few nm}$) were out of the scope of our previous investigations. While we expect that systematic RD measurements for this initial-growth re-

gime will help us understand the mechanism of interface formation, it is also anticipated that the resulting optical data will exhibit complex behaviors since surface and interface layers responsible for the anisotropy are likely to overlap or interact with each other. Another source of complication is the possibility that the dielectric response of such a thin epilayer is probably different from that of bulk.

The ZnSe/GaAs interface is the most critical interface in the blue/green semiconductor lasers.¹⁵⁻¹⁷ Their lifetime, which needs to be extended to more than 1000 h for commercialization, is inversely proportional to the density of the grown-in defects that originate from the ZnSe/GaAs interface.¹⁷ Since the defect generation at this interface is correlated with the structure of the interface,¹⁸⁻²⁰ the mechanism of the interface formation is of technological interest. At the same time, this interface is of scientific interest because it is a good test case for us to study the effects of heterovalency in the epitaxy phenomena.^{9,21-23}

The subject of this study is the RD spectra acquired *in situ* during the initial stages of heteroepitaxial growth of ZnSe on GaAs(001). Combining these results with reflection high-energy electron diffraction (RHEED) observations, we discuss the interface-formation and initial-growth processes. In this materials system, it has been well established that the growth mode can be readily controlled by applying an appropriate pre-growth treatment on the GaAs surface: a Zn treatment on an As-stabilized (2×4) surface for layer-by-layer two-dimensional (2D) growth, and Se-terminated (2×1) for island (3D) growth. This feature allowed us to investigate the effects of the growth mode on the optical properties in a single materials system.

II. EXPERIMENT

The heteroepitaxial growth of ZnSe on GaAs(100) wafers (semi-insulating and nominally on-axis) was carried out in a dual-chamber molecular beam epitaxy (MBE) system. It is equipped with an x-ray photoelectron spectroscopy (XPS)

^{a)}Electronic mail: tyasuda@jrcat.or.jp

^{b)}Also with National Institute for Advanced Interdisciplinary Research (NAIR), Tsukuba 305, Japan.

^{c)}Also with Angstrom Technology Partnership (ATP), Tsukuba 305, Japan.

^{d)}Also with Tsukuba University, Tsukuba, Ibaraki 305, Japan.

^{e)}Also with Institute for Materials Research, Tohoku University, Sendai 980, Japan.

and a scanning tunneling microscope (STM) for on-line surface characterization. Clean GaAs surfaces were first obtained by growing an undoped homoepitaxial layer in the III-V chamber. Prior to ZnSe growth the GaAs surfaces were treated according to two different procedures as identified below:

(a) A Zn exposure on an As-terminated (2×4) GaAs surface. After the exposure the surface remained (2×4),²⁴ and the Zn coverage measured by XPS was less than 1 monolayer.^{25,26} The ZnSe growth on this surface is known to proceed in the 2D mode.²⁰ It is also well established that this procedure is optimal for reducing the defect density in the final ZnSe/GaAs structure.^{18,20}

(b) A Se exposure on a Ga-terminated (4×6) surface followed by annealing at 798 K, which resulted in a Se-terminated (2×1) surface.^{21,24,27} ZnSe growth on this surface proceeded in the 3D mode.^{20,21,24}

In preparing these surfaces, the surface was reconstructed to either (2×4) or (4×6) in the III-V chamber,^{28,29} and the sample was then transferred through a vacuum tunnel to the II-VI chamber where each reconstruction was confirmed by both RHEED and RDS.³⁰ The surfaces were treated by the beam exposure described above for 60 s at 533 K. Special care was taken to prevent uncontrolled adsorption of Se during these procedures. STM observation of the samples that were prepared in separate runs but under the same condition revealed that both surfaces were atomically flat with well-ordered dimer rows.^{24,31,32}

ZnSe growth was carried out in an intermittent manner. The Zn shutter was first opened, and growth was performed by opening the Se shutter for a pre-determined period, typically 4 or 8 s. Each sequence was followed by *in situ* measurements by RDS and RHEED, first under the Zn beam and then under the Se beam. These procedures were repeated up to 44 s. The beam-equivalent pressures of Zn and Se were $(2.3 \pm 0.2) \times 10^{-5}$ and $(4.9 \pm 0.2) \times 10^{-5}$ Pa, respectively. The sample temperature was 533 K throughout the Zn or Se treatment and the ZnSe growth. After a series of the RDS and RHEED measurements were completed, ZnSe growth was continued to a thickness of ~100 nm so the average deposition rate could be determined. We note that the actual deposition rate during the initial growth cycles may be deviated from these average values.

Our home-made RDS setup is similar to the one developed by Aspnes *et al.*,³³ except for the use of a calcite Glan-Taylor prism as the incidence polarizer. The results are displayed in terms of $\Delta\tilde{r}/\tilde{r} = \Delta r/r + i\Delta\theta$, where $\Delta\tilde{r} = r \cdot \exp(i\theta)$ is the complex reflectance. In this article, $\Delta\tilde{r}$ is defined as $\Delta\tilde{r} = \tilde{r}[\bar{1}10] - \tilde{r}[110]$, and only the real-part spectra are shown. The spectra cover an energy range from 1.5 to 5.8 eV.

III. RESULTS

We first present the RDS and RHEED results for the initial stages of ZnSe/GaAs heteroepitaxy. A summary of the observations is given at the end of the section.

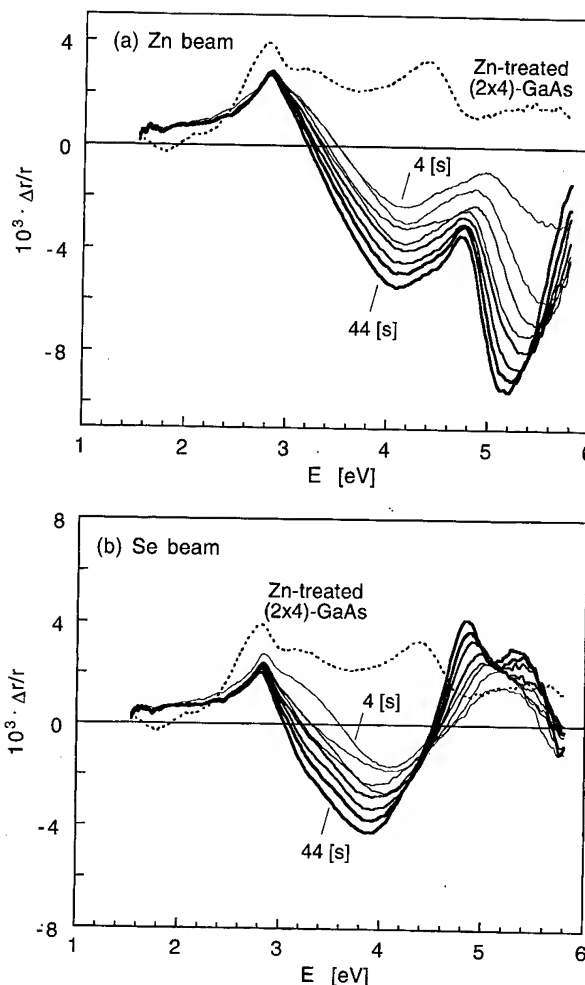


FIG. 1. (a) Evolution of the RD spectra during 2D layer-by-layer growth of ZnSe on a Zn-treated (2×4)-GaAs surface. The spectra were acquired under the Zn beam at 4, 8, 12, 16, 20, 28, 36, and 44 s after initiation of growth. The thickness of the curves is in proportion to the growth period. The estimated deposition rate was 0.10 nm/s. The dotted curve is for the initial Zn-treated (2×4)-GaAs surface. (b) As (a) but under the Se beam.

A. 2D growth mode

Figure 1 shows the RD spectra measured during ZnSe growth on the Zn-treated (2×4)-GaAs surface. The spectra were acquired at 4, 8, 12, 16, 20, 28, 36, and 44 s after initiation of growth, under both (a) Zn and (b) Se beams. If the average growth rate of 0.1 nm/s holds for this thin-film regime, 44 s growth translates to 4.4 nm film growth. In Fig. 1(a), the spectral line shape for the initial Zn-treated surface showed only a slight offset from that of the as-prepared (2×4) surface, as discussed previously.⁹ A large change in the line shape was induced upon the first 4 s growth of ZnSe on the GaAs surface. Successive growth through 44 s changed the spectral intensity above ~3.5 eV, while the peak intensity at 2.8 eV was pinned. We also noticed a red shift in the peak position around 5 eV. The spectra under the Se beam [Fig. 1(b)] also showed a systematic increase in the intensity while the 2.8 eV peak was pinned. An additional feature is a negative dip near 5.1 eV, which became apparent

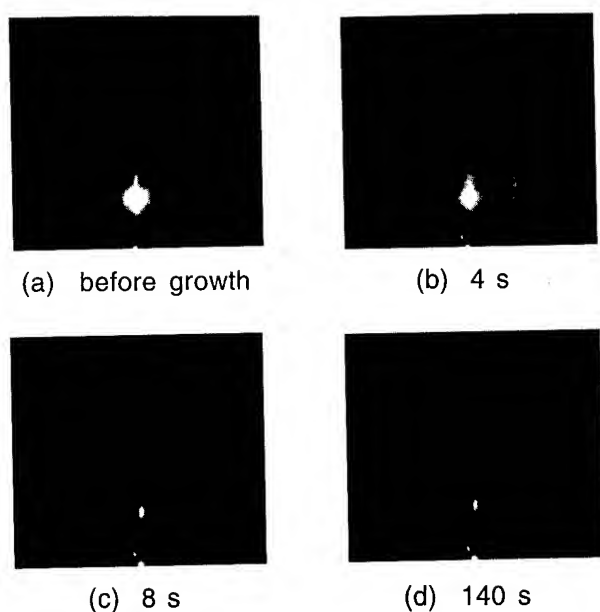


FIG. 2. RHEED patterns during the experiment in Fig. 1. The observation was done under the Se beam and with electron-beam incidence in [110]: (a) (2×4) before growth; (b) (1×1) after 4 s; (c) appearance of (2×1) after 8 s; (d) fully developed (2×1) after 140 s growth.

after 20 s. It is tentatively ascribed to the linear electro-optic (LEO) effect at E_1 and $E_1 + \Delta_1$ transitions of ZnSe.⁴ We note that the surface under the Se or Zn beam was so stable that repeated measurements on the same surface gave exactly the same spectra. Thus an annealing effect at the growth temperature was negligible.

The RHEED patterns for the experiment of Fig. 1 are displayed in Fig. 2. The surface reconstruction of ZnSe is commonly (2×1) for Se termination and $c(2\times 2)$ for Zn termination.³⁴ The patterns in this figure were observed under the Se beam, with the electron-beam incidence set along [110] to detect the fractional diffraction of the (2×1) pattern. The (2×4) pattern of the initial GaAs surface disappeared at 4 s growth, and the (2×1) pattern became visible at 8 s. Observation under the Zn beam with the [010] beam incidence showed that the $c(2\times 2)$ reconstruction was vaguely visible at 12 s. At either incidence direction the patterns were streaky from the beginning through the end of the experiment, confirming that the growth proceeded in the 2D mode.

Figure 3 illustrates the dynamic changes of the RHEED specular-beam intensity and the RD signal at 4.5 eV. A RHEED oscillation was clearly observed during the first and second growth cycles, further evidencing 2D growth. The separation of the two negative peaks in the first cycle, which we assume corresponds to a bilayer of ZnSe, gives a growth rate of 0.12 nm/s, which is close to the average rate of 0.10 nm/s. The dynamic change of the RD signal is consistent with the spectral measurements in Fig. 1. The first 4 s growth induced a decrease of the signal, while for the second growth cycle the ZnSe surface termination changed gradually from Zn to Se during the 4 s period, then the Zn-terminated surface was re-established after the Se shutter was closed.

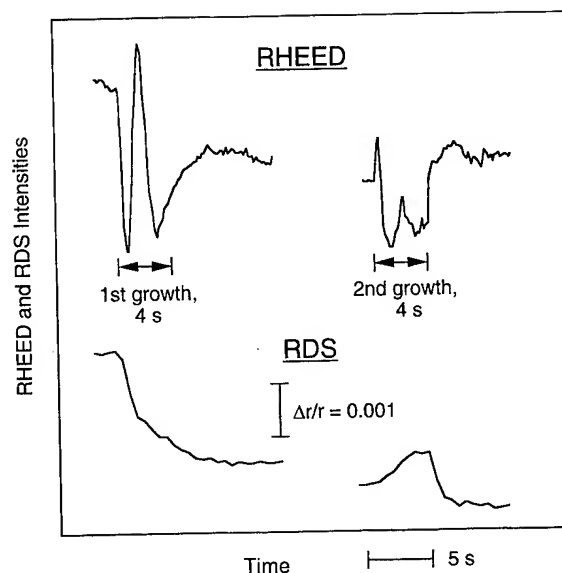


FIG. 3. Dynamic RHEED oscillation (specular beam) and RD response at 4.5 eV for the first and second cycles of intermittent growth (1 cycle=4 s). See Fig. 1 for the RD spectrum after each cycle.

B. 3D growth mode

Figure 4 shows a series of RD spectra taken similarly to those in Fig. 1, but for the 3D growth on the Se-terminated (2×1)-GaAs. From the cross-section transmission electron microscope (TEM) observation described later, we judge the actual growth rate for this data set was again close to the average rate of 0.075 nm/s, even though the growth proceeded through island formation and coalescence. Thus the estimated L_{ZnSe} after 44 s growth is 3.3 nm. The initial GaAs surface exhibited a spectral line shape characteristic of Se-terminated GaAs(001) surfaces, as discussed in detail previously.⁹ The spectral behaviors in Fig. 4 are remarkably different in many aspects from what we saw in Fig. 1:

- (1) the first 4 s growth induced only a slight change in the spectrum;
- (2) the amplitude above 3.5 eV first increased through 16 s, then decreased after 20 s. The spectra at 16 s and 20 s were almost identical to each other;
- (3) switching the beam from Zn to Se induced relatively small changes;
- (4) the red shift of the peak near 5 eV was observed only after 20 s;
- (5) the 2.8 eV peak was pinned through 20 s, but its amplitude started to decrease thereafter;
- (6) the peak intensity at 2.8 eV was much higher than in Fig. 1.

Changes in the RHEED patterns corresponding to the data of Fig. 4 are summarized in Fig. 5. As in Fig. 2, RHEED data were taken under the Se beam and with a [110] probing incidence. A streaky (2×1) pattern for the initial surface turned to a diffuse spotty one after the first 4 s growth, which confirms that the growth mode was 3D. The pattern remained spotty though 44 s, each spot sharpened with in-

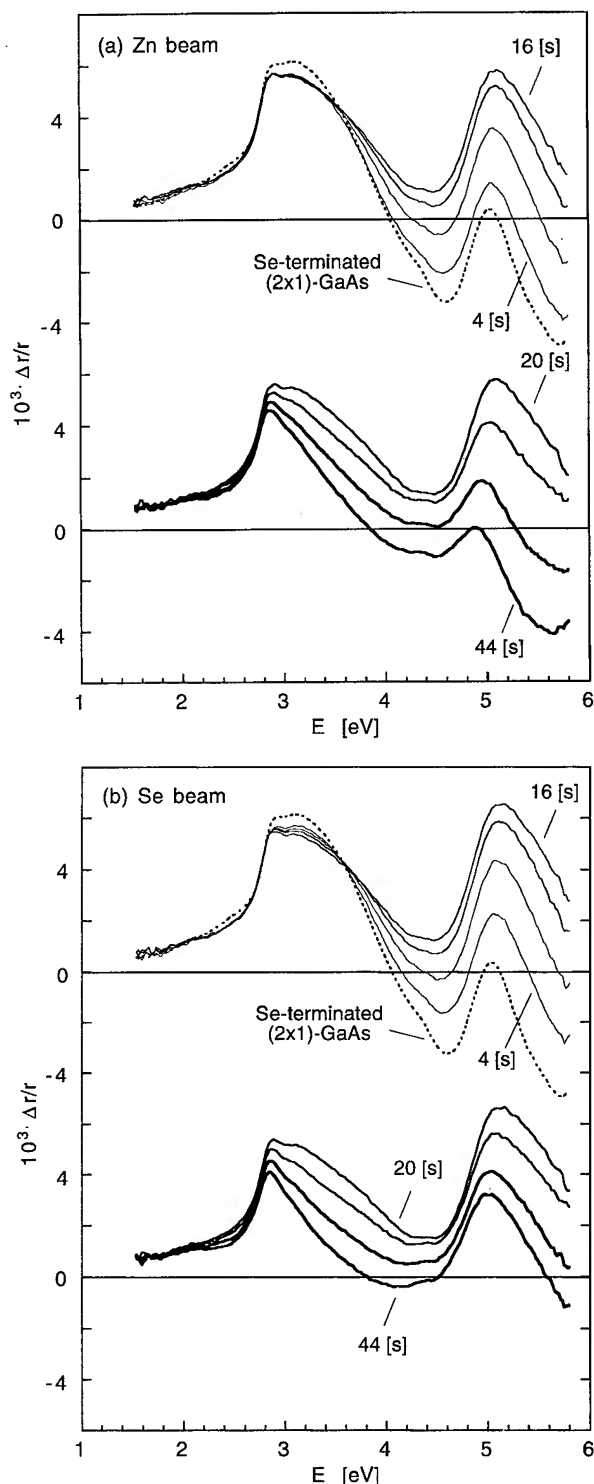


FIG. 4. As in Fig. 1, but for 3D island growth on the Se-terminated (2×1)-GaAs surface, measured under (a) the Zn beam and (b) the Se beam. In both plots, the spectral intensity above 3.5 eV first increased through 16 s, then decreased after 20 s. Spectra after 20 s are shown with a negative offset to make them easier to follow. The estimated deposition rate was 0.075 nm/s. The dotted curve is for the initial Se-terminated (2×1)-GaAs surface.

creasing growth time. The (2×1) pattern appeared at 44 s, but it took ~100 s to obtain a streaky (2×1) pattern. While 16 s was the point when the RD signal around 5 eV took a maximum in Fig. 4, the RHEED data exhibited nothing par-

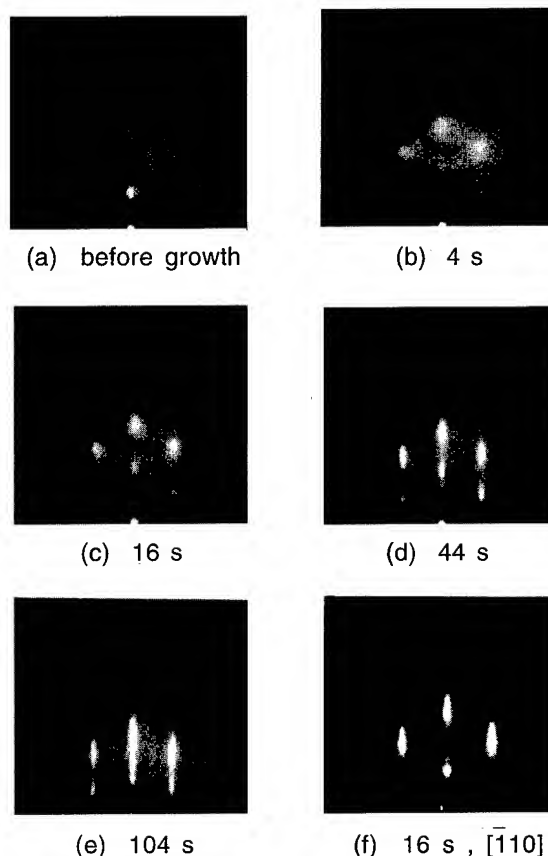


FIG. 5. RHEED patterns during the experiment in Fig. 4. The data were taken under the Se beam. Electron-beam incidence was along [110] except for (f) which was taken with $\bar{1}10$ incidence. (a) (2×1) before growth; (b)–(d) spotty patterns after 4, 16, and 44 s, respectively; (e) a streaky (2×1) pattern after 104 s; (f) spotty but much sharper pattern than (c) after 16 s.

ticularly remarkable at this point. The pattern (f) was taken at 16 s, but with $\bar{1}10$ incidence. Comparing it to (c), the former pattern was much sharper than the latter, indicating that the ZnSe islands were elongated in [110]. This result is consistent with the STM results in the literature, where formation of elongated islands on the Se-terminated GaAs surface was observed in real space.^{21,24}

C. Summary of observations

The behavior of the observed RD spectra for the initial ZnSe growth on GaAs is, as anticipated, fairly complicated. While it is formidable at the present stage to explain all the features we observed, it is possible to extract a few characteristic features that we believe are essential and therefore should be explained.

- (1) The 2.8 eV peak is pinned regardless of ZnSe surface termination and L_{ZnSe} , whereas its amplitude depends on the initial surface.
- (2) For the 2D case, the spectrum above 3.5 eV evolves as growth proceeds. The intensity increases and the peak position near 5 eV exhibits a redshift.

- (3) The behavior for the 3D case is quite different from the 2D case. The intensity takes a maximum at 16 s (or 1.2 nm) under the conditions employed.

In the next section we focus our discussion on these observations.

IV. DISCUSSION

A. Interface-induced peak at 2.8 eV

1. Origin

In our previous studies of the RD spectra of the ZnSe/GaAs structure, we showed that the 2.8 eV peak originates from the interface.^{4,9} This conclusion was deduced from the following observations:

- (i) After separating the measured RD spectrum into surface and interface components using the model calculation, the 2.8 eV peak remained in the interface spectrum but not in the surface spectrum.
- (ii) As seen in Figs. 1 and 4, the position and amplitude of this peak were hardly changed upon switching the surface termination between Zn and Se, while its intensity was dependent on the reconstruction and pre-growth beam treatment of the initial GaAs surface where the interface was formed.

As we pointed out in our previous paper,⁹ the energy position of this 2.8 eV peak is near the E_1 transition of bulk GaAs at the growth temperature of 533 K,³⁵ but it does not agree with critical-point energies of possible interfacial compounds such as Ga_2Se_3 and Zn_3As_2 . This fact suggests that the states associated with this peak can be described as interface-modified states of bulk GaAs. We note that the LEO effect may induce RD features near the E_1 transition energy, however the observed 2.8 eV peak is much broader than what would be expected for the LEO signal, indicating that the contribution of the LEO effect is minor.

Nakayama and Murayama performed *ab initio* calculations to obtain the electronic structure of the ZnS/GaP interface,²³ which is supposedly similar to the ZnSe/GaAs system. They showed that the interface anisotropy near the E_1 transition of GaP is caused by optical transitions between bulk states deformed by charge transfer at the interface. Related to this effect, the surface-induced RD features, which appear shifted relative to the bulk critical points, were recently discussed by Rossow *et al.*³⁶ They identified two types of anisotropy spectra (screening and derivativelike), and pointed out that the derivativelike line shapes can be explained by considering a dynamic localization of bulk-derived electronic states in a semi-infinite system. Further theoretical and experimental investigations are apparently necessary to establish a systematic understanding of surface- and interface-induced RD features located near bulk critical points.

2. Intensity

In Fig. 6 we compare RD spectra for five samples prepared on different initial surfaces and under different growth

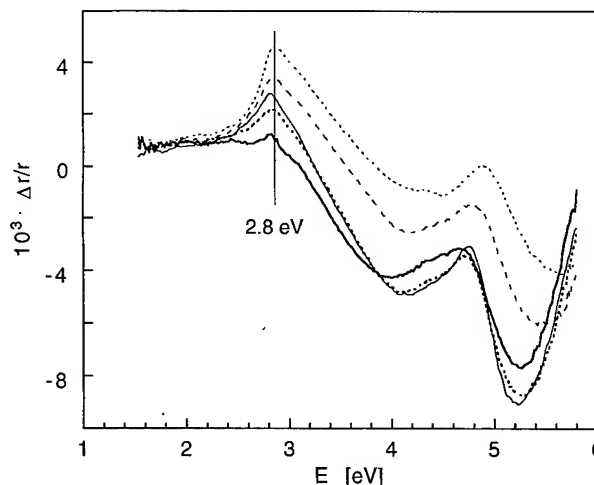


FIG. 6. Effects of the initial-surface preparation and the growth condition on the RD spectra for the ZnSe/GaAs heterostructures. The estimated ZnSe thickness is in the range from 3 to 4 nm for all five samples. The spectra were measured under the Zn beam. Thin dotted: intermittent growth on Se-terminated (2×1) [44 s spectrum in Fig. 4(a)]; broken: alternating beam supply on Zn-treated (2×4); thin solid: intermittent growth on Zn-treated (2×4) [36 s spectrum in Fig. 1(a)]; thick dotted: continuous growth on Zn-treated (2×4) with a lower Se/Zn beam ratio (<2); thick solid: continuous growth on Zn-treated $c(4\times 4)$.

conditions. For all the samples L_{ZnSe} is in the range from 3 to 4 nm. Here we focus our attention on the intensity of the 2.8 eV peak. As will be discussed in the following, we find that the peak intensity is correlated to the Ga–Se bond density at the interface.

In Fig. 6, the sample grown on the Se-terminated (2×1) surface shows the highest 2.8 eV peak, while the sample grown on $c(4\times 4)$ GaAs shows the lowest. For the former case, we showed in our previous study by combining RDS and TEM, that the interfacial layer is composed of Ga and Se and that its thickness is about 4 monolayers.⁹ For the latter case, the same study identified an interfacial layer composed of Zn and As, which indicates that this interface has, if any, only a small density of Ga–Se bonds.

The three curves in between the above two extreme cases were all grown on Zn-treated (2×4), but with different *availability* of Se with respect to Zn. Among these three curves shown, the sample grown by alternating the supply of Zn and Se (i.e., the atomic-layer-epitaxy mode) shows the highest peak. Since the Zn coverage on the beam-treated GaAs surface is less than 1 monolayer,^{25,26} during the first cycle of alternating beam exposure the GaAs surface has an opportunity to react with the Se beam to form Ga–Se bonds. In the intermittent mode, such a Ga–Se reaction is presumably suppressed because Se is consumed to form ZnSe. The sample grown under Se-deficient, i.e., Zn-rich, beam conditions showed the lowest peak among the three.

Thus, the trend seen in Fig. 6 can be well understood if the 2.8 eV peak intensity is correlated with the Ga–Se bond density (or the extent of the Ga–Se reaction) at the interface. An important implication of this result is that the interface stoichiometry is not unique but process-dependent for this

heterostructure. At a heterovalent interface, the valence of each constituent atom should be satisfied through atomic rearrangement, reconstruction, and/or mixing. As estimated theoretically by Nakayama and Oda,²² the interface energy of ZnSe/GaAs is dependent on the interfacial atomic configuration. One might expect that under such a constraint on the valence matching there exists a unique chemical composition of the interfacial layer that minimizes the interface energy. The process-dependent change of the 2.8 eV peak in Fig. 6 implies that kinetics rather than energetics controls the formation of the ZnSe/GaAs interface. Finally we add that the relative availability of Se has a great impact on the defect generation at the ZnSe/GaAs interface as we demonstrated in our recent publications.¹⁸⁻²⁰

3. Pinning

As we have discussed so far, the optical response of the ZnSe/GaAs interface involves two kinds of chemical bonds, i.e., Ga-Se and Zn-As, and the peak intensity at 2.8 eV is a qualitative measure of the fraction of the former. Extending this argument, the peak pinning we observe in Figs. 1 and 4 indicates that once the first several monolayers of ZnSe are deposited, the chemical composition of the interface is fixed throughout the successive growth. In other words, once the interface is formed, atomic intermixing is minimal as far as the optical response of the GaAs side of the interface is concerned. This argument is supported by our annealing experiment of 10 nm ZnSe on Zn-treated (2×4)-GaAs, where the 2.8 eV peak showed little change in intensity nor line shape up to 723 K.

For the 3D case in Fig. 4, the 2.8 eV peak is pinned up to 16 s but starts decreasing after 20 s. This is due to the fact that the surface-induced RD signal near 2.8 eV starts increasing in the negative direction after 20 s. The evolutions of the surface-induced RD signal are discussed in detail in Sec. IV B.

B. Spectral evolution above 3.5 eV

Here we discuss possible factors that contribute to the spectral evolutions observed in Figs. 1 and 4.

1. Development of the surface electronic structure

Figure 7(a) shows the difference spectra calculated from the data set of Fig. 1. They were obtained by subtracting the RD spectrum under Zn from that under Se at the same L_{ZnSe} . Since these differences are induced by switching the surface termination from Zn to Se, the spectra in Fig. 7(a) solely represent the surface-sensitive component of the RD spectrum. In other words they are essentially free from contributions from interface- and bulk-induced anisotropies. In this plot, the peak intensity near 5 eV increases with L_{ZnSe} , and the peak position shows a redshift.

The surface RD signal is approximately expressed by the following linear equation:³

$$\Delta\tilde{r}/\tilde{r} = -(4\pi id/\lambda) \cdot [\Delta\tilde{\epsilon}_0/(\tilde{\epsilon}_s - 1)]. \quad (1)$$

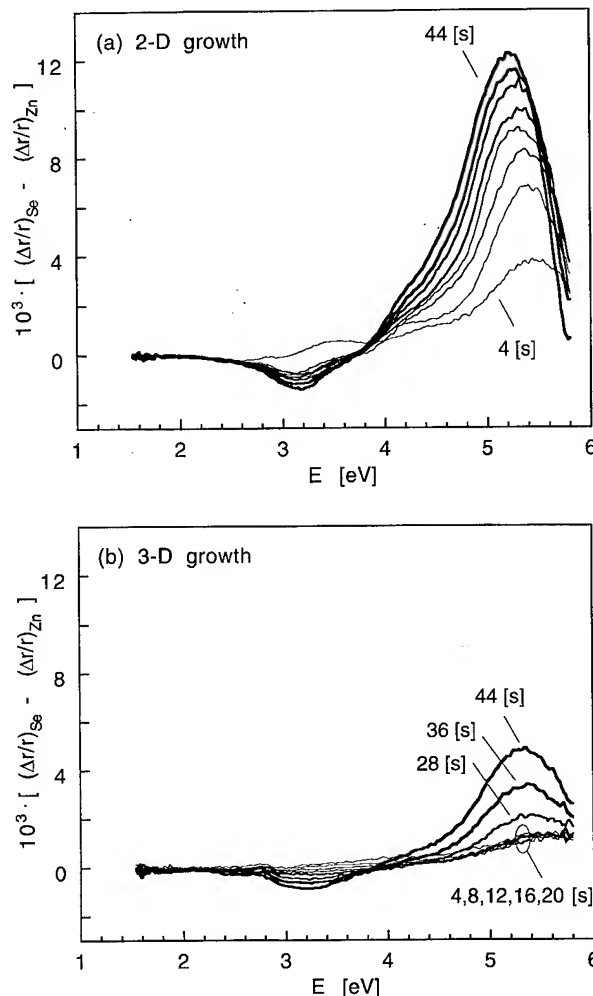


FIG. 7. Evolution of surface-sensitive component in the RD spectra in Figs. 1 and 4. Shown are the difference spectra obtained by subtracting the spectrum measured under the Zn beam from that under the Se beam. The thickness of the curves is in proportion to the growth period: (a) 2D growth on Zn-treated (2×4)GaAs; (b) 3D growth on Se-terminated (2×1)-GaAs.

Here d , $\Delta\tilde{\epsilon}_0$, $\tilde{\epsilon}_s$, and λ are, respectively, the effective thickness and the dielectric anisotropy of the surface layer, the isotropic substrate dielectric function, and the wavelength of light. This equation holds for $d \ll \lambda$. To apply this equation to the present ZnSe/GaAs system, we regard the entire ZnSe/GaAs structure as a virtual substrate, and its pseudo dielectric function is used in place of $\tilde{\epsilon}_s$.³⁷ Then, an increase in RD is either by

- (i) an increase in $\Delta\tilde{\epsilon}_0 \cdot d$, which means a development of the surface electronic structure, or
- (ii) a decrease in the pseudo dielectric function, which indeed takes place as L_{ZnSe} is increased.

While the latter is a purely optical effect, the former is a novel finding if it actually contributes to the observed RD change. Using Eq. (1), we converted the Fig. 7(a) data to $\Delta\tilde{\epsilon}_0 \cdot d$, the result of which is shown for the imaginary part in Fig. 8. It should be remembered that the spectra in this figure represent the changes induced by switching from Zn termination surface to Se termination. The pseudo dielectric func-

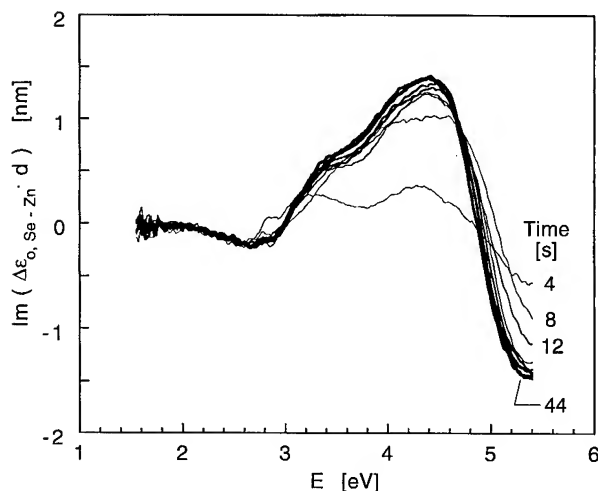


FIG. 8. Evolution of the surface-induced dielectric-anisotropy spectra obtained by converting the Fig. 7(a) data according to the three-phase approach. Shown is the imaginary-part component of $\Delta\epsilon_0 \cdot d$.

tion was calculated by a conventional procedure.³⁸ In the calculation we assumed that d is much smaller than L_{ZnSe} . If this assumption does not hold, the spectra in Fig. 8 are overcorrected for the optical effect. The dielectric functions of GaAs at 500 K and ZnSe at 473 K were taken from literature.^{35,39} Since the experimentally determined ZnSe dielectric functions terminate at 5.1 eV, we extended them to 5.4 eV by referring to the oscillator-model approximation of the ZnSe dielectric functions at room temperature.⁴⁰

From Fig. 8, it is now clear that the initial increase of the surface-RD signal reflects the development of the surface electronic structure. The apparent changes after 16 s (~ 1.6 nm) may be ascribed to the optical effect, however the validity of the assumption employed (i.e., $d \ll L_{\text{ZnSe}}$) needs to be checked before we concur on this point. We point out that the appearance of the surface-induced RD anisotropy from the beginning of the growth is consistent with the RHEED data in Fig. 2, where the surface reconstruction was also visible from the early stages.

A similar evolution of the RD spectra with increasing epilayer thickness was reported for AlAs on GaAs(001) by Wasserman *et al.* In this case, however, the RD spectrum saturated as early as 5 monolayers of AlAs growth (1.4 nm) and no further evolution was observed.⁶ This apparent difference between AlAs/GaAs and ZnSe/GaAs possibly has to do with the fact that AlAs/GaAs is homovalent (III–V on III–V), while ZnSe/GaAs is heterovalent. Since the latter has an interfacial layer with a rather complex structure extending to several monolayers,^{9,20} the surface properties may be affected by the interface particularly when the epilayer is very thin.

2. Effect of the growth mode

Figure 7(b) shows the difference spectra obtained similarly to Fig. 7(a), but for the 3D growth in Fig. 4. In contrast to the 2D case in Fig. 7(a), the difference spectra are featureless with no significant changes until 16 s. It is only after 20



FIG. 9. High-resolution cross-sectional TEM images of coalesced ZnSe islands on Se-terminated GaAs. ZnSe growth was terminated when the RD intensity at 5.2 eV reached a maximum after 16 s growth (see Fig. 4). Horizontal arrowheads indicate the position of the interface.

s that a line shape that looks similar to the 2D case begins to emerge. In Figs. 4(a) and 4(b) the RD signal characteristically changes at 16–20 s. The RHEED data in Fig. 5, however, show nothing particularly remarkable at 16 s.

In order to characterize the structure of the ZnSe film at this point of growth, we performed cross-sectional TEM on a sample that was grown under conditions identical to that of Fig. 4, but for only 16 s. During this short run, we repeated the RD measurement to confirm that the evolution was reproduced. The result of the high-resolution TEM is shown in Fig. 9. We clearly see that at 16 s the islands have just completed coalescing and the ZnSe surface is about to become smooth. L_{ZnSe} determined from Fig. 9 is about 1.4 nm, which agrees with the thickness estimated from an average growth rate of 0.075 nm/s. Thus the evolution of the surface-RD signal for the 3D growth is associated with the formation of a continuous and smooth ZnSe film.

While the local structure of the surface termination is usually a concern for the interpretation of the surface-induced optical anisotropies, the data in Figs. 7(a) and 7(b) indicate that the epilayer thickness and long- or mid-range order in the lateral direction are certainly among the factors to be considered, at least for the initial stages of heteroepitaxial growth.

3. Anisotropic roughness

We finally address the contribution of anisotropic morphology to the initial changes in the Fig. 4 data. As pointed out in Sec. III, the RHEED patterns indicate that the ZnSe islands on the Se-terminated (2×1) surface were elongated in the $[110]$ direction. This result is consistent with STM studies by Li and Pashley concerning ZnSe nucleation on a similarly-prepared (2×1) surface.²¹ Their atomic-resolution image revealed that the islands were indeed elongated in the $[110]$ direction. We also observed a similar roughness-induced anisotropy for the ZnSe film grown in the 3D mode.²⁴

To estimate the contribution of the morphology to the RD signal, we employ the well-established three-phase

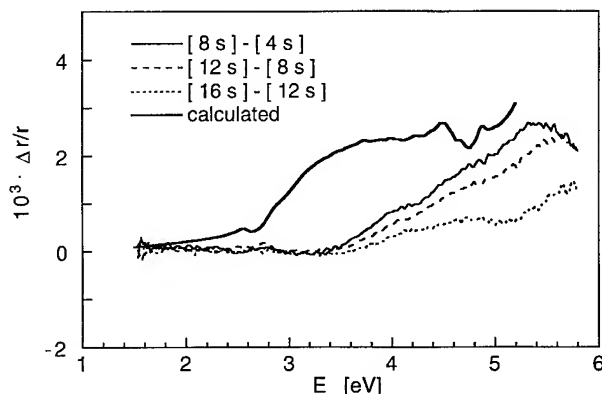


FIG. 10. Thick solid line: simulation of a roughness-induced RD line shape. Thin solid and dashed lines: incremental changes in the first four sequential spectra in Fig. 4(a).

approach,⁴¹ which successfully explained the roughness-induced RD signal for the lattice-mismatched InAs/GaAs system.¹⁰ We consider the case where ZnSe islands are dispersed on the GaAs surface. The dielectric response of such a discontinuous ZnSe layer is approximated by the effective-medium theory.^{2,38} We assume that the thickness and void fraction of this layer to be 1 nm and 0.5, respectively. We also assume the screening factors to be 1 and 5 for $\Gamma 10$ and $\Gamma 110$ polarizations, respectively, which represents islands elongated along $\Gamma 110$.

The result of the simulation is shown in Fig. 10. To check whether a roughness effect can account for the initial increase of the RD signal above 3.5 eV, we also plot in the same figure the increments obtained by subtracting, for instance, the 4 s spectrum from the 8 s one for the Fig. 4(a) series. Although the positive sign of the RD changes can be explained by the optical model described above, the experimental and simulated results do not agree. We note that adjusting the parameters in the effective-medium calculation did not help improve the line shape agreement. Thus, the anisotropic roughness by itself cannot account for the initial RD change in Fig. 4, although it remains one of the possible contributors to the observed changes.

V. CONCLUSIONS

We have carried out systematic RDS and RHEED measurements during the initial stages of ZnSe MBE growth on GaAs(001) surfaces. We compared the results for 2D growth on the Zn-treated (2×4) surface, to those for 3D growth on the Se-terminated (2×1) surface. The interface-induced 2.8 eV peak, which is associated with the E_1 transition of GaAs, is pinned during the initial growth stage, indicating that the interface formation is completed once a ZnSe layer of several monolayers is deposited. The peak intensity at this energy is qualitatively correlated with the Ga–Se bond density at the ZnSe/GaAs interface, which implies that the chemical stoichiometry of the interfacial layer is not unique but clearly process-dependent. The surface-induced signal above 3.5 eV evolves with ZnSe thickness, reflecting the development of

the surface electronic structure. When growth proceeds in the 3D mode, the surface-induced RD features appear only after the surface becomes smooth following coalescence of the islands. Possible contributions of anisotropic morphology to the initial increase of the RD signal were discussed.

ACKNOWLEDGMENTS

This study, partly supported by New Energy Development Organization (NEDO), was performed at JRCAT under a research agreement between NAIR and ATP. The authors would like to thank Professor R. W. Collins at Pennsylvania State University for providing them with the unpublished dielectric-function data.

- ¹D. E. Aspnes, *Thin Solid Films* **233**, 1 (1993).
- ²B. Drevvillon and V. Yakovlev, *Physics of Thin Films*, edited by M. H. Francombe and J. L. Vossen (Academic, San Diego, CA, 1994), Vol. 19, p. 2.
- ³D. W. Aspnes and A. A. Studna, *Phys. Rev. Lett.* **54**, 1956 (1985).
- ⁴T. Yasuda, K. Kimura, S. Miwa, L. H. Kuo, C. G. Jin, K. Tanaka, and T. Yao, *Phys. Rev. Lett.* **77**, 326 (1996).
- ⁵O. Acher, S. M. Koch, F. Omnes, M. Defour, M. Razeghi, and B. Drevvillon, *J. Appl. Phys.* **68**, 3564 (1990).
- ⁶M. Wasserman, I. Kamiya, D. E. Aspnes, L. T. Florez, J. P. Harbison, and P. M. Petroff, *J. Vac. Sci. Technol. B* **9**, 2263 (1991).
- ⁷T. Yasuda, D. E. Aspnes, D. R. Lee, C. H. Bjorkman, and G. Lucovsky, *J. Vac. Sci. Technol. A* **12**, 1152 (1994).
- ⁸Z. Yang, I. K. Sou, Y. H. Yeung, G. K. L. Wong, J. Wang, C. X. Jin, and X. Y. Hou, *J. Vac. Sci. Technol. B* **14**, 2973 (1996).
- ⁹T. Yasuda, L. H. Kuo, K. Kimura, S. Miwa, C. G. Jin, K. Tanaka, and T. Yao, *J. Vac. Sci. Technol. B* **14**, 3052 (1996).
- ¹⁰E. Steimetz, J.-T. Zettler, W. Richter, D. I. Westwood, D. A. Woolf, and Z. Sobiesierski, *J. Vac. Sci. Technol. B* **14**, 3058 (1996).
- ¹¹B. Koopmans, B. Richards, P. V. Santos, and M. Cardona, *23rd International Conference on Physics on Semiconductors, Berlin, July 1996* (World Scientific, Singapore, 1996), p. 1891.
- ¹²J.-T. Zettler, K. Stahrenberg, W. Richter, H. Wenisch, B. Jobst, and D. Hommel, *J. Vac. Sci. Technol. B* **14**, 2757 (1996).
- ¹³C. C. Kim and S. Sivannanthan, *Mater. Res. Soc. Symp. Proc.* **406**, 319 (1996).
- ¹⁴M. J. Kastner, B. Hahn, R. Blumberg, E. Sossna, R. Duschl, and W. Gebhardt, *J. Cryst. Growth* **170**, 188 (1997).
- ¹⁵R. L. Gunshor and A. V. Nurmikko, *Mater. Res. Bull.* **20**, 15 (1995).
- ¹⁶G. F. Neumark, R. M. Park, and J. M. DePuydt, *Physics Today* **47**, 26 (1994).
- ¹⁷A. Ishibashi, *J. Cryst. Growth* **159**, 555 (1996).
- ¹⁸L. H. Kuo, K. Kimura, T. Yasuda, S. Miwa, C. G. Jin, K. Tanaka, and T. Yao, *Appl. Phys. Lett.* **68**, 2413 (1996).
- ¹⁹L. H. Kuo, K. Kimura, S. Miwa, T. Yasuda, C. G. Jin, K. Tanaka, and T. Yao, *Appl. Phys. Lett.* **69**, 140 (1996).
- ²⁰L. H. Kuo, K. Kimura, S. Miwa, T. Yasuda, and T. Yao, *J. Electron. Mater.* **26**, 52 (1997).
- ²¹D. Li and M. D. Pashley, *J. Vac. Sci. Technol. B* **12**, 2547 (1994).
- ²²T. Nakayama, *J. Phys. Soc. Jpn.* **61**, 2458 (1992); K. Oda and T. Nakayama, *Jpn. J. Appl. Phys.* **31**, 2359 (1992).
- ²³T. Nakayama and M. Murayama, *23rd International Conference on Physics on Semiconductors, Berlin, July 1996* (World Scientific, Singapore, 1996), p. 939.
- ²⁴S. Miwa, K. Kimura, T. Yasuda, L. H. Kuo, C. G. Jin, K. Tanaka, and T. Yao, *Appl. Surf. Sci.* **107**, 184 (1996).
- ²⁵S. Heun, J. J. Pagel, S. Rubin, and A. Franciosi, *J. Vac. Sci. Technol. B* **14**, 2980 (1996).
- ²⁶S. Miwa (unpublished).
- ²⁷S. Takatani, T. Kikawa, and M. Nakazawa, *Phys. Rev. B* **45**, 8498 (1992).
- ²⁸T. Hashizume, Q. K. Xue, J. Zhou, A. Ichihmiya, and T. Sakurai, *Phys. Rev. Lett.* **73**, 2208 (1994).
- ²⁹Q. Xue, T. Hashizume, J. M. Zhou, T. Sakata, T. Ohno, and T. Sakurai, *Phys. Rev. Lett.* **74**, 3177 (1995).
- ³⁰I. Kamiya, D. E. Aspnes, L. T. Florez, and J. P. Harbison, *Phys. Rev. B* **46**, 15894 (1992).

- ³¹S. Miwa, L. H. Kuo, K. Kimura, T. Yasuda, and T. Yao, *Jpn. J. Appl. Phys.* **36**, L337 (1997).
- ³²S. Miwa, K. Kimura, T. Yasuda, L. H. Kuo, C. G. Jin, K. Tanaka, and T. Yao, *23rd International Conference on Physics on Semiconductors, Berlin, July 1996* (World Scientific, Singapore, 1996), p. 899.
- ³³D. E. Aspnes, J. P. Harbison, A. A. Studna, and L. T. Florez, *J. Vac. Sci. Technol. A* **6**, 1327 (1988).
- ³⁴T. Yao and T. Takeda, *Appl. Phys. Lett.* **48**, 160 (1986).
- ³⁵P. Lautenschlager, M. Garriga, S. Logothetidis, and M. Cardona, *Phys. Rev. B* **35**, 9174 (1987).
- ³⁶U. Rossow, L. Mantese, and D. E. Aspnes, *23rd International Conference on Physics on Semiconductors, Berlin, July 1996* (World Scientific, Singapore, 1996), p. 831.
- ³⁷D. E. Aspnes, *J. Opt. Soc. Am. A* **10**, 974 (1993).
- ³⁸D. E. Aspnes, *Handbook of Optical Constants of Solids*, edited by E. D. Palik (Academic, San Diego, CA, 1985), p. 89.
- ³⁹R. W. Collins and J. Lee (private communication).
- ⁴⁰S. Adachi and T. Taguchi, *Phys. Rev. B* **43**, 9569 (1991).
- ⁴¹R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light* (Elsevier, Amsterdam, 1977), p. 269.

Explanation of the linear correlation between barrier heights and ideality factors of real metal-semiconductor contacts by laterally nonuniform Schottky barriers

R. F. Schmitsdorf, T. U. Kampen, and W. Mönch^{a)}

Laboratorium für Festkörperphysik, Gerhard-Mercator-Universität Duisburg, D-47048 Duisburg, Germany

(Received 12 January 1997; accepted 1 March 1997)

A new and simple-to-use method to obtain homogeneous Schottky barrier heights from effective barrier heights and ideality factors that are determined from current-voltage (I - V) characteristics of metal-semiconductor contacts is presented. This approach is justified by a theory of metal-semiconductor interfaces with laterally inhomogeneous distributions of barrier heights. Effective barrier heights and ideality factors were determined from I - V characteristics of Si and GaN Schottky contacts and a linear reduction of the effective barrier heights with increasing ideality factors was always observed. These findings are explained by numerical simulations of inhomogeneous Schottky contacts which are based on theoretical results by Tung [Phys. Rev. B **45**, 13509 (1992)]. The homogeneous barrier heights of metal-semiconductor contacts are obtained by a linear extrapolation of the effective barrier heights to $n_{if} \cong 1.01$, the value of the ideality factor characteristic for image-force lowering of Schottky barriers only. © 1997 American Vacuum Society. [S0734-211X(97)07804-9]

I. INTRODUCTION

Most metal-semiconductor contacts are rectifying.¹ Schottky² explained this behavior by depletion layers on the semiconductor side of such interfaces. The band bending in this space-charge region is characterized by its barrier height, which is the energy distance between the Fermi level and the edge of the respective majority-carrier band right at the interface. For moderate doping levels of the semiconductor, tunneling through the barrier may be neglected and thermal emission over the barrier determines the current-voltage (I - V) characteristics of rectifying metal-semiconductor or Schottky contacts. The image force lowers the barrier height and makes it voltage-dependent. This Schottky effect³ is accounted for by an ideality factor. Traditionally, rectifying metal-semiconductor contacts are characterized by their barrier heights and ideality factors.

In the past, discussions on the physical mechanisms that determine the barrier heights in Schottky contacts have dominated the field of metal-semiconductor contacts. The metal-induced gap states⁴⁻⁶ are considered to be the primary mechanism⁷ while additional interface dipoles due to interface doping⁸ or correlated with specific interface structures,⁹⁻¹² as well as fabrication-induced defects,⁷ were proposed as secondary mechanisms. Mostly, the interfaces were implicitly assumed to be laterally uniform and the contacts were characterized by their barrier heights and ideality factors. However, Tung *et al.*^{13,14} and Rau *et al.*¹⁵ already pointed out that inhomogeneities may play an important role and have to be considered in the evaluation of experimental I - V characteristics. The application of standard procedures gives *effective* barrier heights and ideality factors only. Both parameters vary from diode to diode even if they are identi-

cally prepared. Only recently, a correlation between effective barrier heights and ideality factors was reported and it was approximated by a linear relationship.^{8,11} This finding was attributed to inhomogeneous interfaces and the barrier heights obtained by extrapolation to the ideality factor calculated for image-force lowering only were taken as values characteristic of homogeneous interfaces. It was concluded that these values, rather than mean values obtained from a set of identically prepared contacts of the same kind, should be compared with theoretical results. The purpose of the present article is to analyze this procedure by considering theoretical results obtained by Tung¹⁴ for the current-voltage relationship of nonuniform Schottky contacts. It is worth mentioning that effective barrier heights and ideality factors reported for Pd₂Si/ n -Si (Ref. 16) and Ni/ n -GaAs (Ref. 17) Schottky contacts are also linearly correlated. These data were obtained from I - V characteristics recorded as function of temperature.

II. THEORETICAL BACKGROUND

The thermionic-emission theory gives the current across uniform metal-semiconductor interface as^{18,19}

$$I = AA_R^{**} T^2 \exp\left(-\frac{\Phi_{B0}^{eff}}{k_B T}\right) \left[\exp\left(\frac{e_0 V_c}{n k_B T}\right) - 1 \right], \quad (1)$$

where A is the diode area, A_R^{**} is the effective Richardson constant, T is the temperature, k_B is Boltzmann's constant, e_0 is the electronic charge, V_c is the voltage drop across the contact, and Φ_{B0}^{eff} and n are the effective barrier height and the ideality factor of the contact, respectively. Image-force lowering and generation-recombination currents give ideality factors $n_{if} = 1.01 \div 1.03$ and $n_{rg} = 2$, respectively. These mechanisms are not sufficient to explain the experimental values of $1.01 \leq n \leq 2$ routinely observed. Moreover, this

^{a)}Electronic mail: w.moench@uni-duisburg.de

simple thermionic theory fails to account for other anomalies such as, for example, temperature-dependent barrier heights and ideality factors. These problems were overcome by a theoretical approach of Tung's^{13,14} who considered nonuniform or "patchy" Schottky contacts.

Initially, inhomogeneous metal-semiconductor contacts were treated as to consist of separate diodes with different barrier heights and areas in parallel. Such models are correct as long as the dimensions of the patches are large compared to the Debye length of the semiconductor. If the size of the patches embedded in much larger areas of higher and uniform barrier height becomes comparable to or even smaller than the Debye length, then saddle-point barriers exist in front of the patches. The barrier height at the saddle point is intermediate between the values of the patch itself and of the surrounding homogeneous contact area. The saddle-point barrier depends on the applied voltage. This results in a "pinch-off" of the patches as the bias increases.

Tung¹³ has analyzed the current transport in nonuniform Schottky contacts in great detail. His results explain all the "anomalies" of the barrier height and ideality factor which have been obtained from I - V characteristics of metal-semiconductor contacts when they are evaluated by applying Eq. (1). Tung found that, for example, circular inhomogeneities are characterized by the parameter $\gamma = 3(R_p^2 \Delta_p / 4)^{1/3}$, i.e., by the product of the patch area πR_p^2 and the deviation Δ_p of their local barrier height from the homogeneous value Φ_{B0} . The subscript 0 refers to zero bias. Tung then assumed patches of area density ρ_p with a Gaussian distribution

$$N(\gamma) = \frac{\rho_p}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\gamma^2}{\sigma^2}\right) \quad (2)$$

of their patch parameters γ , where σ is the standard deviation. The total current through such patchy diodes then results as

$$I_{\text{total}} = AA_R^{**} T^2 \exp\left(-\frac{\Phi_{B0}}{k_B T}\right) \left[\exp\left(\frac{e_0 V_c}{k_B T}\right) - 1 \right] \times \left[1 + \frac{8\pi\rho_p\sigma^2\eta^{1/3}}{9(V_{b0} - V_c)^{1/3}} \exp\left(\frac{e_0^2\sigma^2(V_{b0} - V_c)^{2/3}}{2(k_B T)^2\eta^{2/3}}\right) \right] \quad (3)$$

with $\eta = \epsilon_b \epsilon_0 / e_0 N_d$ where ϵ_b and N_d are the bulk dielectric constant and the dopant density of the semiconductor, respectively, and V_{b0} is the interface band-bending of the uniform barrier outside the patches.

In real Schottky contacts the series resistance R_s of the semiconductor bulk and the measurement setup cannot be omitted.²⁰ This reduces the applied voltage V_a by $R_s I_{\text{total}}$ and the voltage across the contact itself amounts to $V_c = V_a - R_s I_{\text{total}}$. Equation (3) then becomes an implicit function of the total current and has to be rewritten as

$$I_{\text{total}} = AA_R^{**} T^2 \exp\left(-\frac{\Phi_{B0}}{k_B T}\right) \left[\exp\left(\frac{e_0(V_a - R_s I_{\text{total}})}{k_B T}\right) - 1 \right] \times \left[1 + \frac{8\pi\rho_p\sigma^2\eta^{1/3}}{9(V_{b0} - V_a + R_s I_{\text{total}})^{1/3}} \exp\left(\frac{e_0^2\sigma^2(V_{b0} - V_a + R_s I_{\text{total}})^{2/3}}{2(k_B T)^2\eta^{2/3}}\right) \right] \quad (4)$$

Equation (4) completely describes the current through nonuniform Schottky contacts which exhibit circular patches with a Gaussian distribution of the patch parameter γ . However, Eqs. (3) and (4) do not include the image-force lowering of the barrier heights.

III. EXPERIMENT

As substrates we used standard n -Si(111) wafers (Wacker Chemitronik) and n -type α -GaN epilayers on sapphire (Cree Research Inc.). Both types of samples were first dipped in hydrofluoric acid which was diluted by a buffered HF solution (HF:NH₄F:NH₄OH) with $pH=9$. This procedure leaves the Si samples with hydrophobic Si(111):H-1 \times 1 surfaces. The Si substrates were subsequently transferred into the ultrahigh vacuum (UHV) system and briefly annealed at 850 °C to desorb the hydrogen and to produce clean Si(111)-7 \times 7 surfaces. After the wet chemical treatment the GaN samples were rinsed in de-ionized water, blown dry with N₂ gas, and immediately transferred into the UHV system. The GaN substrates were still contaminated by ~ 1 monolayer of oxygen. It was completely removed during exposure of the GaN samples to a flux of 1×10^{16} Ga atoms/cm² s at 800 °C for 10 min, which was followed by an annealing in UHV for another 30 min. These surfaces showed sharp 1 \times 1 low-energy electron diffraction (LEED) patterns and their x-ray photon spectroscopy (XPS) spectra revealed no traces of any residual impurities.

Onto the clean Si(111)-7 \times 7 and GaN(0001)-1 \times 1 surfaces, metals were evaporated *in situ* from Knudsen cells through a mask which had circular openings with nominal diameters of 1 mm. The exact diode areas were determined using an optical microscope. The evaporation rates were monitored using a quartz oscillator and were adjusted to 0.05 nm/s. The metal films had a nominal thickness of 150 nm. Ohmic contacts were achieved by rubbing metallic gallium with an Al pencil onto the GaN samples. The Si wafers, on the other hand, exhibited implanted n^+ layers on their rear face which acted as Ohmic contacts. The I - V characteristics were recorded outside the UHV system and in the dark.

IV. RESULTS AND DISCUSSION

Figure 1 displays the forward I - V characteristic (open circles) of a Sn/Si(111) Schottky contact and a least-squares fit (full line) of Eq. (4) to the experimental data. The fitting parameters were the homogeneous barrier height Φ_{B0} , the series resistance R_s , the patch density ρ_p , and the standard deviation σ of the patch-parameter γ . The diode temperature was 296 K. The fitting procedure included all experimental

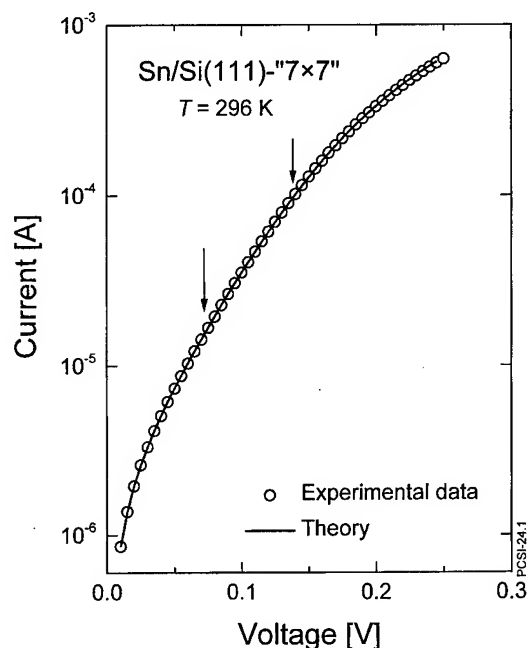


FIG. 1. Current-voltage characteristic of a Sn/Si(111)-"7×7" Schottky diode at room temperature. The full line is a least-squares fit of Eq. (4) to the experimental data with $\Phi_{B0}=0.679$ eV, $\sigma=2.18\times10^{-4}$ cm^{2/3} V^{1/3}, $\rho_p=1.79\times10^7$ cm⁻², and $R_s=97$ Ω. The arrows indicate the voltage range from which an effective barrier height $\Phi_{B0}^{\text{eff}}=0.62$ eV and an ideality factor $n=1.41$ were calculated using Eq. (1).

data points in the voltage range from 0.01 to 0.25 V. We obtained a homogeneous barrier height $\Phi_{B0}=0.679$ eV, a series resistance $R_s=97$ Ω, a patch density $\rho_p=1.79\times10^7$ cm⁻², and a standard deviation of the patch-parameter $\sigma=2.18\times10^{-4}$ cm^{2/3} V^{1/3}. In addition to this fitting procedure we also performed a "standard evaluation" of the I - V characteristics using Eq. (1). In the voltage range from 0.075 to 0.14 V, which is indicated by the arrows in Fig. 1, $\ln(I)$ varies linearly as a function of the applied voltage. According to Eq. (1), the respective slope parameter gives the ideality factor and the extrapolation to zero applied voltage gives the zero-bias barrier height. With the effective Richardson constant $A_R^{**}=112$ A cm⁻² K⁻² (Ref. 21) and the diode area $A=7.85\times10^{-3}$ cm² we obtained an effective barrier height $\Phi_{B0}^{\text{eff}}=0.62$ eV and an ideality factor $n=1.41$. The homogeneous and the effective barrier heights of this Sn/Si(111) contact thus differ by $\Phi_{B0}-\Phi_{B0}^{\text{eff}}=0.059$ eV.

Effective Schottky barrier heights and ideality factors vary from diode to diode, therefore, it is common practice to take averages. Figure 2 displays histograms of the effective barrier heights and ideality factors for as many as 38 Sn/Si(111) Schottky contacts. The statistical analysis yields the mean effective barrier height $\langle\Phi_{B0}^{\text{eff}}\rangle=(0.64\pm0.017)$ eV and the mean ideality factor $\langle n\rangle=1.22\pm0.09$. Such mean values disregard the pronounced correlation between effective barrier heights and ideality factors.

Figure 3 now displays the effective barrier heights of the same 38 diodes as a function of their ideality factors. We would like to emphasize that the barrier heights plotted in

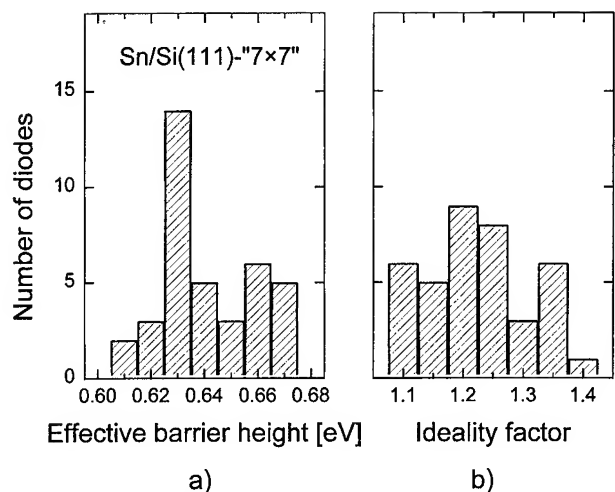


FIG. 2. Effective Schottky barrier heights (a) and ideality factors (b) obtained from I - V characteristics of 38 Sn/Si(111)-"7×7" diodes. The mean values of the effective barrier heights and of the ideality factors are $\langle\Phi_{B0}^{\text{eff}}\rangle=0.64$ eV and $\langle n\rangle=1.22$, respectively.

Figs. 2 and 3 result from the standard evaluation of the experimental I - V curves using Eq. (1) rather than from fitting procedures employing Eq. (4). Figure 3 clearly confirms the well-known linear correlation^{8,11} between effective barrier heights and ideality factors. In general, large effective barrier heights are related with small and then close to unity ideality factors. This trend may be understood in terms of nonuniform interfaces.

Both the ideality factor and the effective barrier heights depend, in a complicated way, on the patch density and the standard deviation of the patch-parameter. Due to the implicit form of Eq. (4), there is no analytic expression available. Therefore, we performed a numerical simulation of the effective barrier height and the ideality factor as a function of the patch density for different standard deviations of the patch-parameter. The calculation procedure was as follows. With the parameters chosen we first calculated I - V curves using Eq. (4). We then applied the standard evaluation pro-

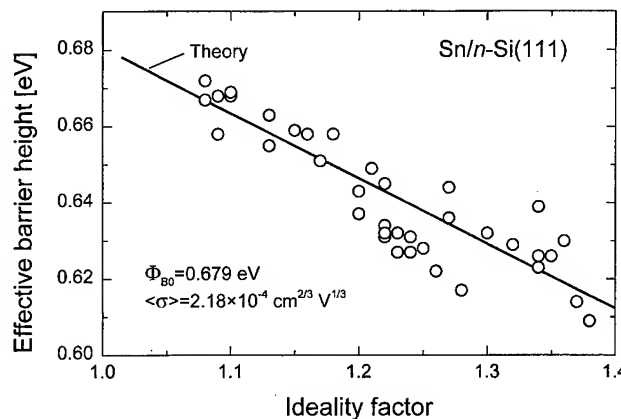


FIG. 3. Effective barrier heights of Sn/Si(111)-"7×7" diodes as a function of their ideality factors at room temperature. The data are the same as in Fig. 2. The full line is the result of a numerical simulation.

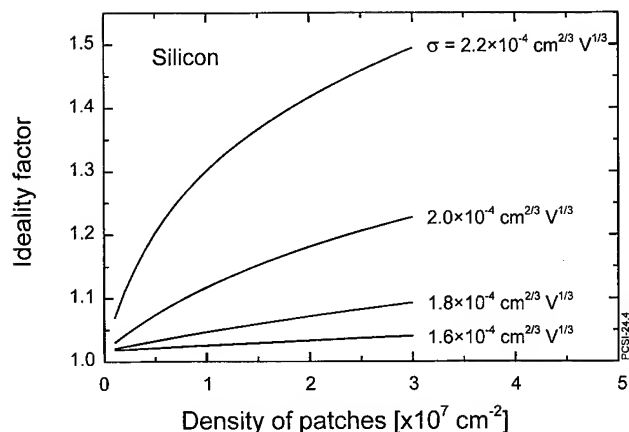


FIG. 4. Simulation of the ideality factor as a function of the patch density for four different standard deviations σ of the patch parameter γ . The silicon contact was assumed to have a homogeneous barrier height $\Phi_{B0}=0.679$ eV and a series resistance $R_s=80$ Ω and the doping density of the substrate was set to $N_d=1.4 \times 10^{15}$ cm^{-3} .

cedure using Eq. (1) to these calculated I - V curves and obtained effective barrier heights and ideality factors. Figure 4 displays our results for the ideality factors. As one might intuitively expect, the larger the patch density and/or the standard deviation of the patch-parameter, the larger the respective ideality factor. Therefore, the ideality factors represent a direct measure of the interface uniformity. This finding and the assumption that the patches have smaller barrier heights than the homogeneous contact explains the experimentally observed reduction of the barrier heights with increasing ideality factors.

The full line in Fig. 3 shows the effective barrier as a function of the ideality factor which were both obtained from our simulations as described in the preceding paragraph. The parameters used were a homogeneous barrier height $\Phi_{B0}=0.679$ eV, a series resistance $R_s=90$ Ω , and a standard variation of the patch parameter $\sigma=2.18 \times 10^{-4}$ $\text{cm}^{2/3}$ $\text{V}^{1/3}$. This choice was motivated by the parameters obtained from the fit of Eq. (4) to the experimental data displayed in Fig. 1. The patch density ρ_p increased up to about 2×10^7 cm^{-2} . The result of our simulations is indeed a straight line. It is worth mentioning that a "simple" linear regression of the experimental data yields the same straight line and, as a consequence, the same homogeneous barrier height. This finding confirms our earlier conclusion that homogeneous barrier heights may be obtained by linear extrapolations of Φ_{B0}^{eff} vs n plots to $n=n_{if}=1.01 \div 1.03$, the image-force-controlled ideality factor of homogeneous Schottky contacts. This procedure is much easier than the determination of the homogeneous barrier heights by fitting the complicated Eq. (4) to many experimental I - V curves.

In our numerical simulations we made the simplifying assumption that all diodes are characterized by one and the same standard deviation σ of the patch-parameter. Actually, the standard deviation of the patch-parameter may vary from diode to diode. This might explain the scatter of the experimental data around the straight line in Fig. 3.

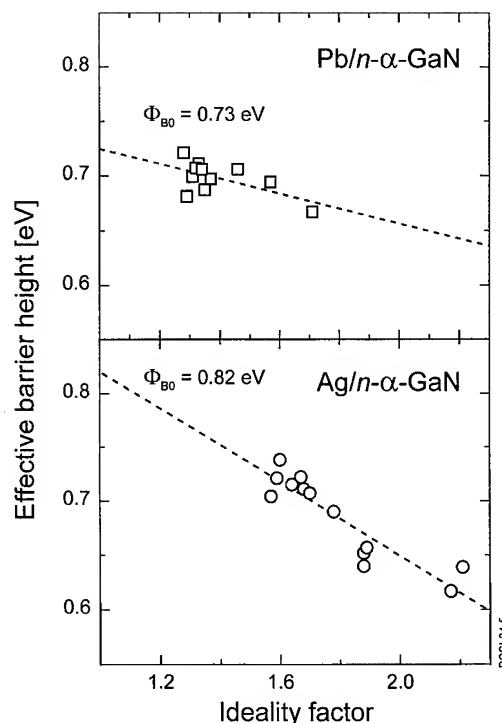


FIG. 5. Effective barrier heights of Pb- and Ag/ n - α -GaN diodes as a function of their ideality factors at room temperature. The dashed lines are linear least-squares fits to the experimental data.

Linear correlations between effective Schottky barrier heights and ideality factors were also observed for a number of other metal-semiconductor combinations.^{8,11} As examples, results of Ag/Si(111) and Ag as well as Pb/GaN(0001) Schottky contacts are displayed in Figs. 5 and 6. The data of the GaN Schottky contacts clearly show the importance of Φ_{B0}^{eff} vs n plots. Figure 5 gives the Schottky barrier heights of homogeneous Pb and Ag/GaN contacts as 0.73 and 0.82 eV, respectively. The simple statistical analysis of the data, on the other hand, yields *mean* effective barrier heights of 0.70 and 0.69 eV, respectively. The mean effective and the homogeneous Schottky barrier heights of the Ag/GaN contacts

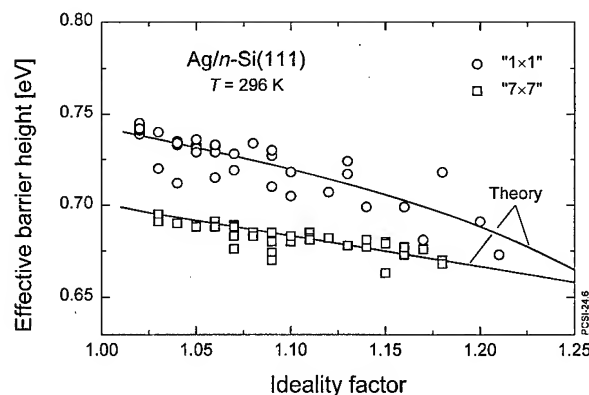


FIG. 6. Effective barrier heights of Ag/Si(111)-"7 \times 7" and -1 \times 1 diodes as a function of their respective ideality factors at room temperature. The full lines are results of numerical simulations.

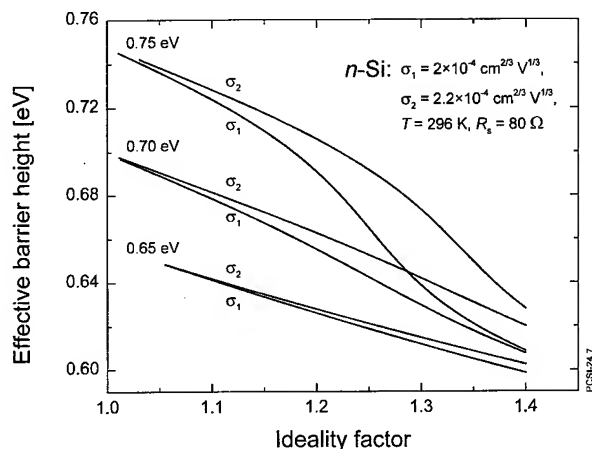


FIG. 7. Simulation of the effective barrier height as a function of the ideality factor for three different homogeneous barrier heights and two different standard deviations each of the patch parameter. The silicon substrate was assumed to have a doping density $N_d = 1.4 \times 10^{15} \text{ cm}^{-3}$ and a series resistance $R_s = 80 \text{ } \Omega$.

differ by as much as 130 meV. This large discrepancy is obviously caused by the large ideality factors of the Ag/GaN diodes. They might be a consequence of the poor quality of the GaN epilayers used.

Effective Schottky barrier heights will always decrease as a function of the ideality factors. However, whether the correlation is linear depends on the actual inhomogeneities. A best fit to the Ag/Si(111)-“7×7” data of Fig. 6 is achieved with $\Phi_{B0} = 0.70 \text{ eV}$ and $\sigma = 2.35 \times 10^{-4} \text{ cm}^{2/3} \text{ V}^{1/3}$ and gives a strict linear correlation. The Ag/Si(111)-“1×1” data, on the other hand, are best described by the parameters $\Phi_{B0} = 0.745 \text{ eV}$ and $\sigma = 2.012 \times 10^{-4} \text{ cm}^{2/3} \text{ V}^{1/3}$. The fit is linear up to $n \approx 1.15$ and is slightly bent above this value. Such nonlinearities will be discussed by considering further results of the numerical simulations outlined above.

Figure 7 shows Φ_{B0}^{eff} vs n plots for three different homogeneous Schottky barrier heights and two different standard deviations of the patch parameter each. The $\Phi_{B0}^{\text{eff}}(n)$ curves become nonlinear above $n \approx 1.2$ for large effective barrier heights and small standard deviations of the patch-parameter. The curves may be divided into three regions. First, the correlation is close to linear for ideality factors near to unity. In the second region, the curve exhibits a strong downward bent. It is eventually followed by a third and almost linear part for large ideality factors. In this third region the patch density becomes large. Then, the effective barrier height of the contact is almost completely controlled by the patches and the homogeneous Schottky barrier height is of less importance. Therefore, some care must be taken if experimental data are analyzed in terms of $\Phi_{B0}^{\text{eff}}(n)$ plots and the ideality factors are large. In such cases a detailed analysis of individual I - V characteristics using Eq. (4) should be carried out. However, this procedure requires the input of material parameters such as the modified Richardson constant, the exact doping density, the dielectric constant and the density of states in the majority-carrier band of the semiconductor. If these parameters are not available—as in the case of GaN—

then the extrapolation of the effective barrier heights to $n = n_{if} \approx 1$ should be used as a first-order approximation. At present, we cannot derive an analytic relation for the ideality-factor range in which Φ_{B0}^{eff} vs n plots are linear. This needs further investigation.

V. CONCLUSIONS

It is common practice to determine effective Schottky barrier heights and ideality factors from I - V characteristics of real metal-semiconductor contacts using the I - V relationship derived for thermionic transport across laterally uniform interfaces. Both effective barrier heights and ideality factors vary from diode to diode. However, we found pronounced correlations between the effective barrier heights and the ideality factors and attributed this behavior to local deviations of the barrier height from the value characteristic of uniform interfaces. The homogeneous Schottky barrier height of metal-semiconductor contacts may be easily obtained from plots of the effective barrier heights of a number of contacts versus their ideality factors. If such plots may be fitted by straight lines, then the extrapolation to the ideality factor $n = n_{if} \approx 1$ directly gives the homogeneous Schottky barrier height. We justify this procedure by numerical simulations of I - V curves which use Tung's theory of laterally inhomogeneous contacts with Gaussian distributions of the parameter characterizing such patchy metal-semiconductor interfaces. These homogeneous barrier heights rather than effective barrier heights of individual contacts or mean values should be used to discuss theories on the physical mechanisms that determine the barrier heights of metal-semiconductor contacts. Provided the semiconductor substrate is well characterized then the homogeneous Schottky barrier height may be obtained from the I - V characteristic of even one contact.

ACKNOWLEDGMENT

This work was supported by Grant No. Mo318/16-1 of the Deutsche Forschungsgemeinschaft.

- ¹F. Braun, Pogg. Ann. Phys. **153**, 556 (1874).
- ²W. Schottky, Naturwissenschaften **26**, 843 (1938).
- ³W. Schottky, Phys. Z. **15**, 872 (1914).
- ⁴V. Heine, Phys. Rev. A **138**, 1689 (1965).
- ⁵S. G. Louie and M. L. Cohen, Phys. Rev. B **13**, 2461 (1976).
- ⁶E. Louie, F. Yndurain, and F. Flores, Phys. Rev. B **13**, 4408 (1976).
- ⁷W. Mönch, Phys. Rev. Lett. **58**, 1260 (1987).
- ⁸T. U. Kampen and W. Mönch, Surf. Sci. **331–333**, 490 (1995).
- ⁹R. Tung, Phys. Rev. Lett. **51**, 461 (1984).
- ¹⁰D. R. Heslinga, H. H. Weitering, D. P. van der Werf, T. M. Klapwijk, and T. Hibma, Phys. Rev. Lett. **64**, 1589 (1990).
- ¹¹R. F. Schmitsdorf, T. U. Kampen, and W. Mönch, Surf. Sci. **324**, 249 (1995).
- ¹²R. F. Schmitsdorf, C. Gregor, and W. Mönch (unpublished).
- ¹³R. T. Tung, Phys. Rev. B **45**, 13509 (1992).
- ¹⁴J. P. Sullivan, R. T. Tung, M. R. Pinto, and W. R. Graham, J. Appl. Phys. **70**, 7403 (1991).
- ¹⁵U. Rau, H. H. Güttler, and J. H. Werner, Mater. Res. Soc. Symp. Proc. **260**, 245 (1992).
- ¹⁶S. Chand and J. Kumar, Appl. Phys. A **63**, 171 (1996).
- ¹⁷W. Nathan, Z. Shoeshine, G. Ashkinazi, B. Meyler, and O. Zolotarevski, Solid-State Electron. **39**, 1457 (1996).

¹⁸S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, NY, 1981).

¹⁹E. H. Rhoderick and R. H. Williams, *Metal-Semiconductor Contacts*, 2nd ed. (Clarendon, Oxford, 1988).

²⁰For an overview on the influence of the series resistance on I - V

characteristics see, for example, V. Aubry and F. Meyer, *J. Appl. Phys.* **76**, 7973 (1994).

²¹M. P. Lepselter and S. M. Sze, *Bell Syst. Tech. J.* **47**, 195 (1968).

²²T. U. Kampen, R. F. Schmitsdorf, and W. Mönch, *Appl. Phys. A* **60**, 391 (1995).

Evolution mechanism of nearly pinning-free platinum/*n*-type indium phosphide interface with a high Schottky barrier height by *in situ* electrochemical process

Hideki Hasegawa,^{a)} Taketomo Sato, and Tamotsu Hashizume
Research Center for Interface Quantum Electronics and Graduate School of Electronics
and Information Engineering, Hokkaido University, Sapporo 060, Japan

(Received 13 January 1997; accepted 23 March 1997)

Recently, unusually high Schottky barrier heights (SBHs) have been realized by our group on *n*-type InP and related materials by an *in situ* electrochemical deposition of Pt. In an attempt to understand the underlying mechanism of the SBH enhancement, this article investigates in detail the evolution process of the metal (Pt, Ni, Co, and Ag)-InP interface during the *in situ* electrochemical process, using current-voltage, capacitance voltage, deep level transient spectroscopy, x-ray photoelectron spectroscopy, Raman, atomic force microscope, and scanning electron microscope measurements. Pt deposition by the electrochemical process realized an oxide-free, defect-free, stress-free, and nearly pinning-free interface, whereas Pt deposition by conventional electron beam evaporation and sputtering processes as well as Ag, Ni, and Co deposition by the electrochemical process gave rise to stressed and pinned interfaces. The observed large process dependence of SBH can be explained by none of the metal induced gap state model, the unified defect model, and the effective workfunction model. It is explained here by the disorder induced gap state model.

© 1997 American Vacuum Society. [S0734-211X(97)09404-3]

I. INTRODUCTION

Schottky barrier heights (SBHs) for electrons at metal-semiconductor interfaces of GaAs and related materials such as $\text{Al}_x\text{Ga}_{1-x}\text{As}$ are high, being larger than 900 meV. This has led to successful realization of GaAs metal-semiconductor field-effect-transistors (MESFETs) and high electron mobility transistors (HEMTs). On the other hand, those of the so-called InP-based materials such as InP, $\text{In}_x\text{Ga}_{1-x}\text{As}$ with x being in the vicinity of 0.53 and $\text{In}_x\text{Al}_{1-x}\text{As}$ with x being in the vicinity of 0.52 are generally known to be low. Typical values reported in the literature are in the range of 400–500 meV for InP, 200–300 meV for $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$, and 600–700 meV for $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$, respectively, being almost independent on the metal work function due to the so-called Fermi level pinning phenomenon.

Technologically, these materials are extremely important because they can produce electronic devices with superb high-frequency performances and well-behaved and highly reliable optoelectronic devices matched to wavelength in the long wavelength optical fiber communication. These InP-based devices are expected to become the “key devices” of various advanced systems, in various millimeter-wave wireless local-area and world-wide network systems as well as the fiber-to-the-home (FTTH) information infrastructure. However, low SBH values in these materials are problematic and impose severe limitations on realizability, performance, and reliability of advanced devices such as MESFETs, HEMTs, metal-semiconductor-metal (M-S-M) interdigital photodetectors, etc.

Theoretically, the mechanism for Fermi-level pinning, i.e., how the contact potential difference at the metal-

semiconductor interface is almost completely screened so as to produce metal-independent SBH values, has been a long-standing issue, and various models^{1–7} have been proposed, including the unified defect model (UDM),¹ the metal induced gap state (MIGS) model,^{2,3} the disorder induced gap state (DIGS) model,^{4–6} and the effective workfunction (EWF) model.⁷ Among them, the MIGS model is theoretically most well defined and is perhaps most predominantly accepted currently among physicists. Indeed, this model predicts low SBH values for InP-based materials that are consistent with the above mentioned experimental values. Since the MIGS model asserts that Fermi level pinning is an intrinsic property of the ideal intimate metal-semiconductor interface, it may seem impossible to change them.

However, we have recently shown^{8,9} experimentally that SBH values can be substantially increased by 300–450 meV on *n*-type InP, $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ and $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ by depositing platinum with a novel *in situ* electrochemical process. Using this technique, a well behaved InP MESFET has been realized¹⁰ for the first time which showed even enhancement mode operation due to a high SBH value of 0.86 eV.¹⁰

The purpose of this article is to study in detail the evolution process of the metal-InP interface during the *in situ* electrochemical process in an attempt to understand the underlying mechanism of the observed SBH enhancement. Detailed current-voltage (I - V), capacitance voltage (C - V), deep level transient spectroscopy (DLTS), x-ray photoelectron spectroscopy (XPS), atomic force microscope (AFM), scanning electron microscope (SEM), and Raman studies have been made for Pt, Ni, Co, and Ag/InP interfaces prepared by the electrochemical process and for those by the conventional metal deposition processes such as vacuum evaporation and sputtering.

^{a)}Electronic mail: hasegawa@ryouko.rciqe.hokudai.ac.jp

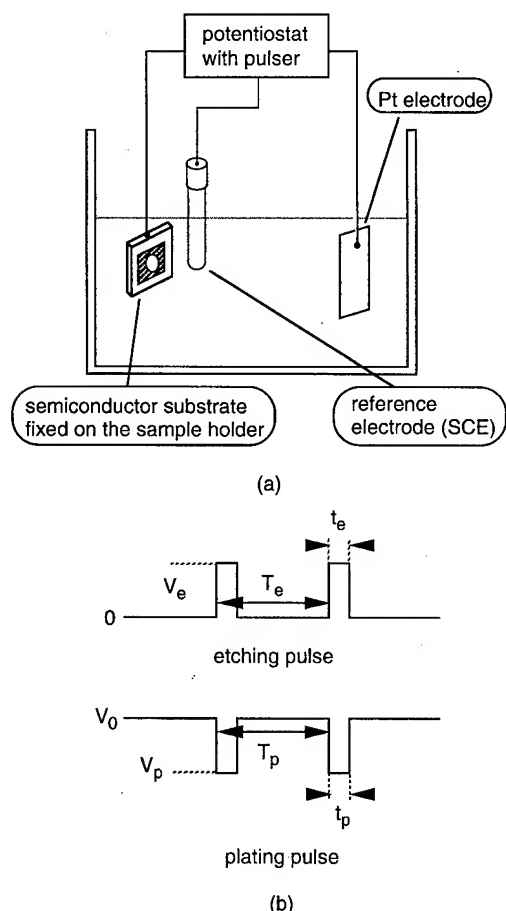


FIG. 1. (a) Set up of the novel *in situ* electrochemical process and (b) pulse waveforms for etching and plating modes.

It has been found that Pt deposition by the electrochemical process realizes an oxide-free, defect-free, stress-free, and nearly pinning-free interface, whereas Pt deposition by the electron beam (EB) evaporation and sputtering processes as well as Ag, Ni, and Co deposition by the electrochemical process produces stressed and pinned interfaces. The observed large processing dependence of SBH cannot be explained by the MIGS model. The observed SBH increase is explained here by the disorder induced gap state (DIGS) model in terms of DIGS density reduction.⁴⁻⁶

II. EXPERIMENT

A. Interface formation

The *in situ* electrochemical process was done using the setup schematically shown in Fig. 1(a). The electrolyte bath contains three electrodes, i.e., a semiconductor electrode onto which metal is deposited, a Pt counterelectrode, and a reference saturated calomel electrode (SCE). The overpotential of the semiconductor electrode with respect to the SCE reference was controlled by a potentiostat.

The process consists of controlled anodic etching of semiconductors followed by subsequent *in situ* cathodic deposition of metal in the same electrolyte realized by the polarity change of potential supply. Both etching and deposition were

done in pulsed modes, using the pulse waveforms shown in Fig. 1(b). In the etching mode, holes are produced near the surface of the semiconductor by avalanche pulses with a height V_e , a width t_e , and a period T_e . These holes oxidize the semiconductor surface where oxides are immediately dissolved into the electrolyte resulting in the so-called anodic dissolution. Typical values of V_e , t_e , and T_e used in this study were 15 V, 1 μ s, and 400 μ s, respectively. After the etching of the surface, plating pulses with a height V_p , a width t_p , and a period T_p superposed on an offset potential of V_0 are immediately supplied to start deposition of metal in the same electrolyte. The values of V_p , t_p , T_p , and V_0 were changed over wide ranges and their effects on the electrical properties and the surface morphology of the plated surface were studied and empirically optimized. The optimal values used in this study were $V_p = 500\text{--}1500$ mV, $t_p = 0.4\text{--}40$ μ s, $T_p = 400$ μ s, and $V_0 = 200\text{--}500$ mV, respectively. The use of pulses instead of direct currents allowed accurate control of the average thickness for etching and metal deposition in sub-nm range. The pulsed mode also enhances the current efficiency for metal plating, largely avoiding the problem of H_2 evolution at the surface. This type of electrochemical process was previously applied to Pt deposition to GaAs¹¹ and GaAlAs/GaAs two-dimensional electron gas (2DEG) systems.¹²

In this study, Pt, Ni, Co, and Ag were deposited onto *n*-type InP, using the following electrolytes:

Pt: 1M HCl(200 ml)+ H_2PtCl_6 + NH_4OH [$pH=1$]

Co: 1M HCl(200 ml)+ $CoSO_4$ (10 g)+ NH_4OH [$pH=1$]

Ni: 1M HCl(200 ml)+ $NiSO_4$ (24 g)+ $NiCl_2$ (3 g)
+ NH_4OH [$pH=2$]

Ag: 1M HCl(200 ml)+AgCl(1 g)[$pH=0$].

For the purpose of comparison, Pt was deposited onto *n*-type InP by the conventional EB evaporation and rf sputtering processes. The beam energy of 10 kV was used for the EB process. For the sputtering process, the rf frequency and the power were 13.56 MHz and 60 W, respectively. The Pt deposition rate was 20 nm/min for both processes.

B. Interface characterization

The electrical properties of the Schottky diodes were investigated by I - V , C - V , and DLTS measurements. To determine SBH from the I - V characteristics, the effective Richardson constant of $9.6\text{ A cm}^{-2}\text{ K}^{-2}$ was used for *n*-type InP. In addition to the conventional $\log J$ (current density) versus V plot, the so-called Rhoderick plot¹³ and Norde plot¹⁴ were made to accurately determine the values of SBH.

The surface morphology of the etched and plated surfaces of InP were investigated by the AFM and SEM, using Nano-scope II and Hitachi S-4100, respectively.

The chemical composition and profile of the interface was studied by x-ray photoelectron spectroscopy (XPS) with Ar-ion sputtering, using Perkin Elmer 5100C spectrometer. The excitation source was Mg $K\alpha$ ($h\nu = 1253.6$ eV). Raman

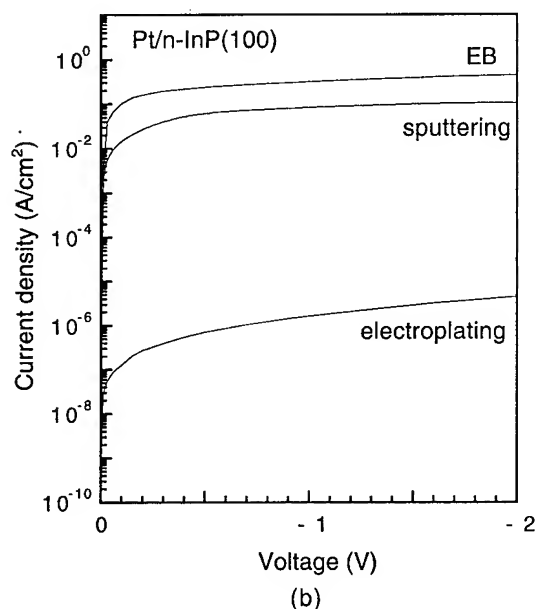
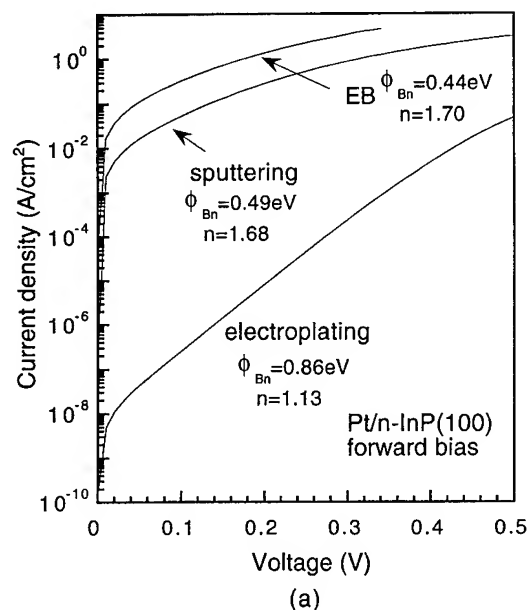


FIG. 2. The measured (a) forward and (b) reverse I - V characteristics of the Pt/ n -InP Schottky diodes produced by the electrochemical process, the conventional EB deposition, and the sputtering process.

spectroscopy measurements of the plated surfaces were made using Jobin Yvon TA 64000 Raman spectrometer.

III. EXPERIMENTAL RESULTS

A. Electrical properties

Examples of the measured forward and reverse I - V characteristics of the Pt Schottky diodes produced by the electrochemical process, the conventional EB deposition, and the sputtering process are compared in Figs. 2(a) and 2(b), respectively. As seen in Fig. 2(a), the latter conventional two process yielded poor I - V characteristics with conventional low values of n -type SBH, ϕ_{Bn} . On the other hand, the elec-

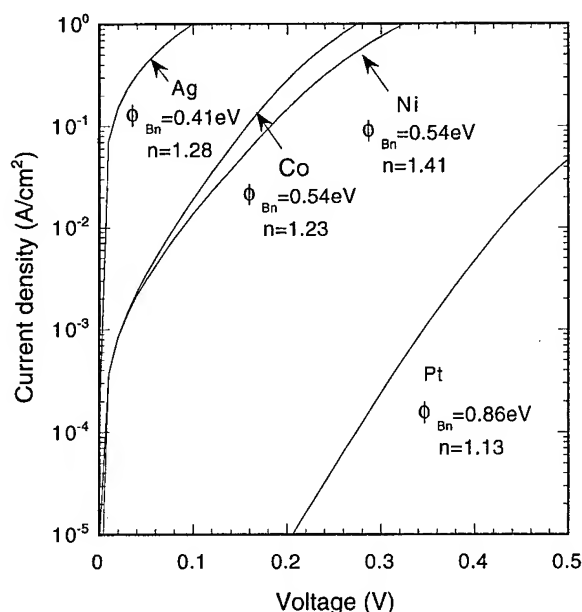


FIG. 3. Forward I - V characteristics of various Schottky diodes for n -InP produced by the *in situ* electrochemical process.

trochemically produced Pt Schottky diodes showed nearly ideal thermionic emission characteristics with high SBHs. As seen in Fig. 2(b), the electrochemical process reduced the reverse leakage currents by a factor as large as 4 to 5 orders of magnitude. Figure 3 shows examples of the forward I - V characteristics of the electrochemically produced n -InP Schottky diodes having different metals. The result showed very much metal-dependent behavior. The highest SBH values were obtained by Pt deposition.

The average values of SBH and ideality factor n determined from the I - V characteristics are summarized in Table I for various Schottky diodes studied here. In addition to conventional $\log J$ - V plots, Rhoderick plots¹³ and Norde plots¹⁴ were made. All these gave fairly consistent SBH values. As seen in Table I, only the Pt/InP diodes prepared by the electrochemical process exhibited much enhanced SBH values of 0.86–0.90 eV with respect to Ni, Co, and Ag di-

TABLE I. Summary of electrical measurements on metal/ n -InP Schottky diodes.

Metal	Process	$\log J$ - V	Rhoderick plot	Norde plot
Pt	electroplating	SBH=0.86 eV n = 1.15	SBH=0.85 eV n = 1.14	SBH=0.90 eV
	EB deposition	SBH=0.44 eV n = 1.70	SBH=0.45 eV n = 1.68	SBH=0.45 eV
	sputtering	SBH=0.49 eV n = 1.68	SBH=0.49 eV n = 1.66	SBH=0.51 eV
Ni	electroplating	SBH=0.47 eV n = 1.32	SBH=0.47 eV n = 1.31	SBH=0.49 eV
Co	electroplating	SBH=0.57 eV n = 1.54	SBH=0.57 eV n = 1.40	SBH=0.63 eV
Ag	electroplating	SBH=0.42 eV n = 1.47	SBH=0.42 eV n = 2.03	SBH=0.41 eV

odes by the electrochemical process and Pt diodes by EB deposition and sputtering all of which gave conventional SBH values of 0.4–0.5 eV.

C–V measurements were also made on Pt/InP diodes formed by the electrochemical process in order to see the possible existence of an interfacial layer or near surface band modification which can give rise to high SBH values. Very straight $1/C^2$ –V plots were obtained whose slopes were consistent with the electron concentration determined by Hall measurements. In all of the electrochemical Pt/InP diodes, excellent agreements of SBH values were obtained between I –V and C–V methods with the difference in SBH values being within 20 meV as already reported elsewhere.⁹ This result strongly indicates that the novel electrochemical process realizes intimate Schottky contacts without an interfacial layer and near-surface modification of the band profile.

No appreciable deep levels were detected by DLTS study on electrochemically prepared Pt/InP Schottky diodes.

B. XPS and Raman study

In order further to check the presence or absence of an interfacial layer on the Pt/InP Schottky diodes prepared by the electrochemical process, an XPS in-depth profile analysis was made, using Ar ion sputtering on various Pt Schottky diodes. The $\text{In}_{3d_{5/2}}$ and P_{2p} spectra observed at the interface are shown in Fig. 4 for electrodeposited, EB-deposited, and sputter-deposited Pt/InP interfaces. The spectra as obtained from an molecular beam epitaxy (MBE) prepared InP free surface subjected to Ar-ion sputtering are also shown on the top as a reference.

As seen in Fig. 4, the spectra from the Pt-electroplated interface remain almost completely the same with those from the Ar-sputtered reference free surface. In contrast to this, EB-deposited and Pt-sputter-deposited interfaces possess spectra that are significantly different from the reference, showing broadening, peak shifts, and inclusion of oxides. Thus, the Pt-electroplated interface forms an intimate contact without inclusion of appreciable interfacial oxide layer.

An additional remarkable feature of the Pt-electroplated interface became apparent by Raman spectra measurements. The Raman spectra obtained from the Pt-electroplated surface and from the Pt-EB deposited surface are shown for various thicknesses of deposition up to about 100 Å in Figs. 5(a) and 5(b), respectively. Those from the Ni-, Co-, and Ag-electroplated surfaces having a metal thickness of about 100 Å are shown in Fig. 6. It is seen from Figs. 5 and 6 that no appreciable Raman peak shift of InP LO phonon peak and broadening takes place during electrodeposition of Pt, whereas large peak shifts and broadening takes place in other cases. Since Raman peak shifts and broadening are related to the presence of stress, the result here shows that Pt-electrodeposited surface is remarkably free from stress.

C. AFM/SEM study

AFM/SEM observation of etched and metal-deposited surfaces was made. Examples of the AFM images are shown in Figs. 7(a)–7(c). They were taken after anodic etching

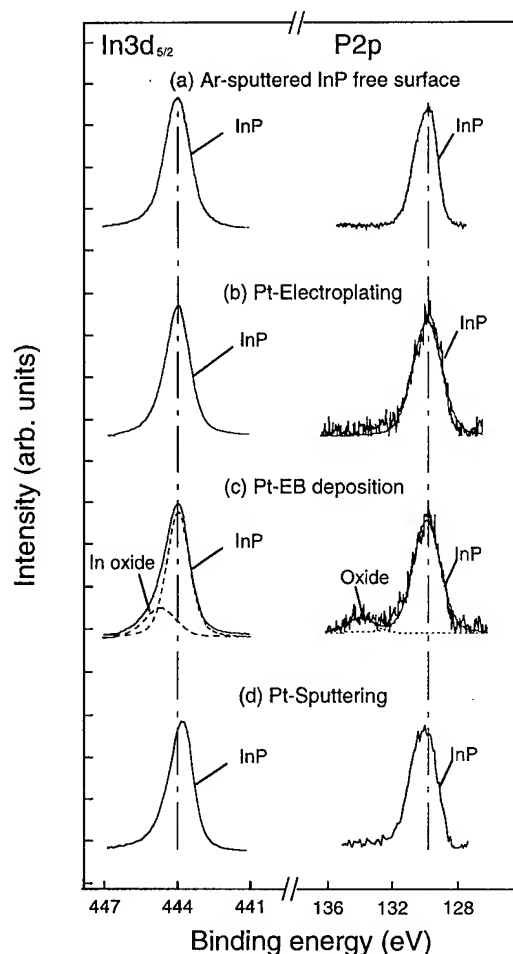


Fig. 4. XPS spectra from (a) the Ar-sputtered InP free surface and the Pt/InP interfaces by (b) the *in situ* electrochemical process, (c) the EB process, and (d) the sputtering process.

[Fig. 7(a)], after cathodic deposition of Pt [Fig. 7(b)], and after EB deposition of Pt [Fig. 7(c)], respectively. As seen in Figs. 7(a) and 7(b), the etched and Pt-electroplated surfaces exhibited a relatively smooth overall topography. The root mean square (rms) roughness of the etched surface was 1.0 nm and this was found to be slightly better than that of the surface etched by conventional wet chemical etchant. The Pt-electroplated surfaces had an rms roughness of 2.4 nm with a dense array of small grains. On the other hand, EB-deposited surface was much more rough with a rms roughness of 7.9 nm and the presence of much larger grains.

The SEM images taken on the initial phases of Pt electroplating at various plating times are shown in Figs. 8(a)–8(c). Deposition starts from formation of small Pt semispherical grains whose sizes are 30–50 nm as seen in Fig. 8(a). Calculations showed that such grains were formed at an average rate of 100–300 grains per pulse depending on the details of the pulse waveforms. Longer plating times did not increase the grain size but increased the number of such grains, as shown in Fig. 8(b). The grains eventually became dense enough and produced a smooth surface as shown in Fig. 8(c).

On the other hand, the surface morphology was found to be very different for Ni-, Co-, and Ag-electroplated surfaces

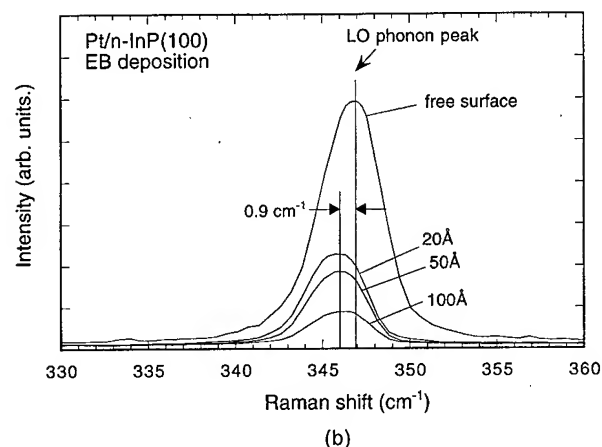
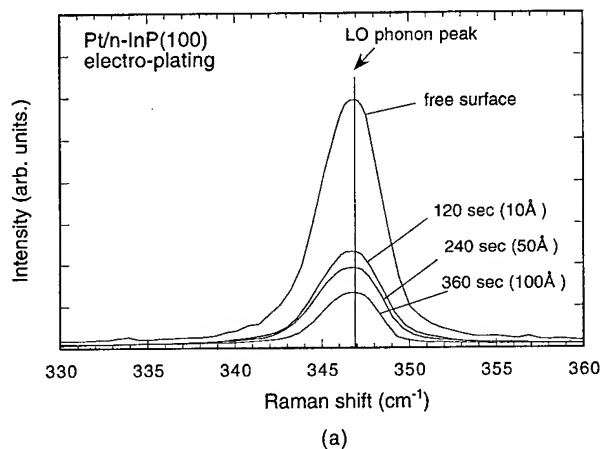


FIG. 5. Raman spectra in backscattering geometry from the Pt/InP interfaces by (a) the *in situ* electrochemical process and (b) the EB process.

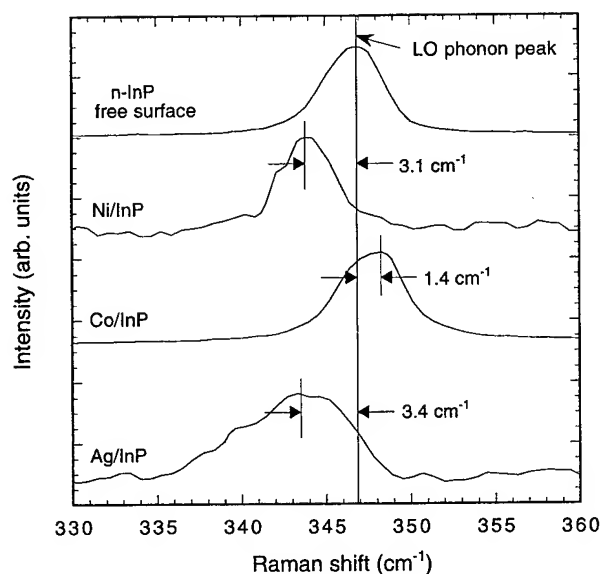


FIG. 6. Raman spectra in backscattering geometry from the metal/InP interfaces by the *in situ* electrochemical process.

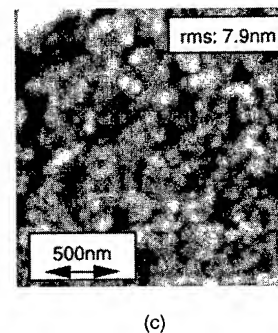
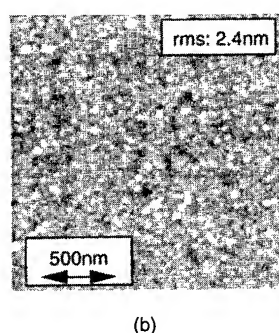
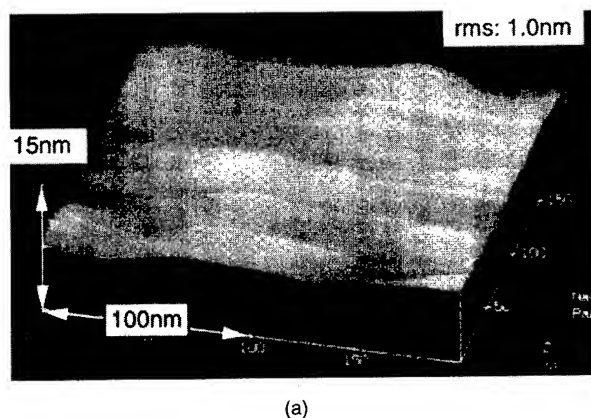


FIG. 7. AFM images of (a) anodic etched InP surface and the Pt-deposited InP surfaces by (b) the *in situ* electrochemical process and (c) the EB process.

as shown in Figs. 9(a)–9(c), respectively. Namely, grains sizes are much larger, and grains tend to form larger islands. These features can also be seen in the EB deposited Pt surface shown in Fig. 7(c). It is likely that these larger grains and island formations give random stress to the InP surface, resulting in strain induced Raman shifts. However, such features are completely absent in the case of the electrodeposition of Pt, and this appears to be the reason for the stress-free interface, although it is not clear at present which property of the metal or the electrodeposition process is responsible for this observed difference among metal species.

IV. DISCUSSION

The observed SBH values of *n*-type InP Schottky diodes studied here are plotted in Fig. 10 versus the metal work function by Michaelson.¹⁵ Our previous data⁸ taken on the Schottky diodes prepared by vacuum deposition are also included. It is noted in Fig. 10 that, except for the Pt by electrodeposition, most of the data fall within the range of 400–500 meV without showing clear dependence on the metal work function. The average value of the so-called slope factor S , defined by $d\phi_{Bn}/d\phi_m$ with ϕ_{Bn} and ϕ_m being the *n*-type SBH and the metal work function, respectively, is about 0.15 from the least square fitting, as shown by the dotted line in Fig. 10. Thus, only the electroplated Pt con-

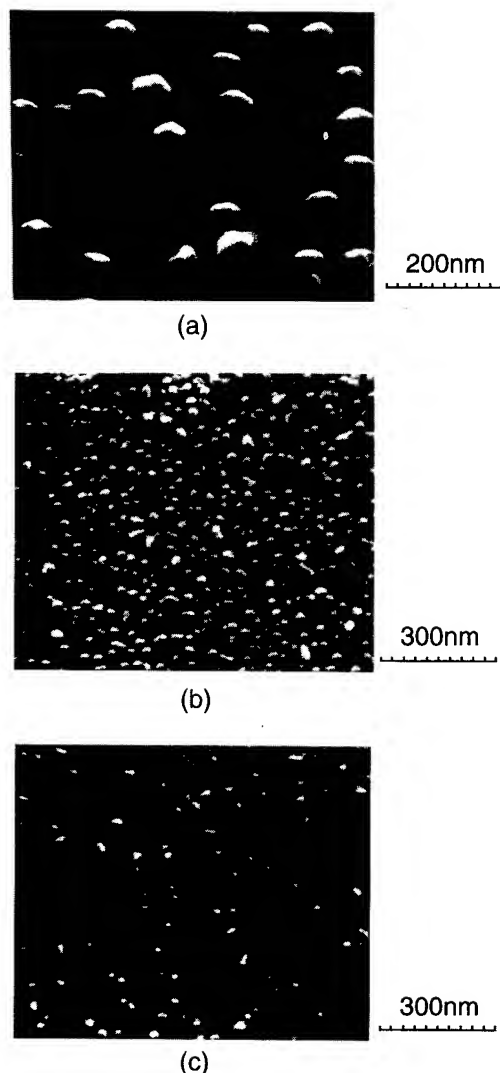


FIG. 8. SEM images of initial phases of Pt deposition on *n*-InP for deposition time of (a) 180, (b) 240, and (c) 360 s by the *in situ* electrochemical process.

tacts show remarkably enhanced SBH values of near 0.9 eV, while others show the conventional pinned value with some data scatters.

As mentioned previously, the currently predominant theoretical model for Fermi level pinning at an intimate metal contact is the MIGS model.^{2,3} According to this model, the SBH for an *n*-type semiconductor is given in the infinitely strong screening limit by

$$\phi_{Bn} = E_C - E_{CNL}, \quad (1)$$

where E_{CNL} is the charge neutrality level of the MIGS continuum. As for the theoretical location of E_{CNL} , one may use the calculated midgap energy of $E_V + 0.76$ eV after Tersoff¹⁶ or that of the dielectric midgap energy of $E_V + 0.87$ eV by Cardona and Christensen,¹⁷ giving rise to $\phi_{Bn} = 0.59$ and 0.48 eV, respectively. These values, particularly the latter, are in good agreement with the experimentally observed

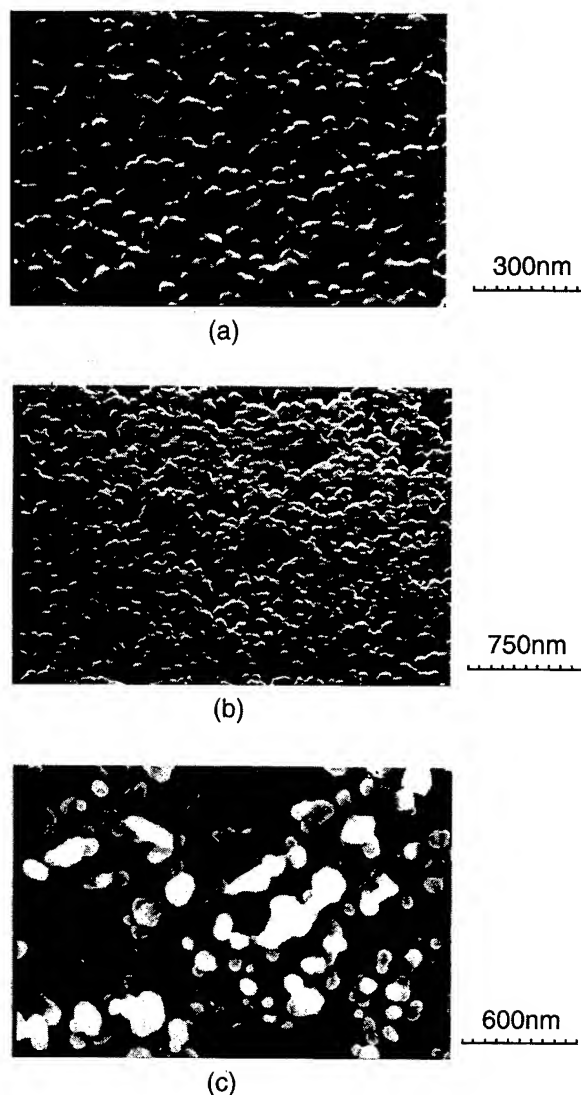


FIG. 9. SEM images of (a) Ni-, (b) Co-, and (c) Ag-deposited InP surfaces by the *in situ* electrochemical process.

SBH values of 400–500 meV for various diodes except the Pt-electroplated diodes, but are far away from those of the Pt-electroplated diodes by 300–400 meV.

Of course, one has to take account of the fact that the screening capability of the MIGS continuum against the contact potential difference is not infinite. This effect may be taken account of by the following formula by Mönch.¹⁸

$$\phi_{Bn} = (E_C - E_{CNL}) + S(\phi_m - \phi_{CNL}) \quad (2)$$

with

$$S = \frac{1}{1 + \frac{q^2}{\epsilon_i \epsilon_0} D_{\text{MIGS}}(E_{CNL}) \delta_i}, \quad (3)$$

where ϕ_m is the metal work function, ϕ_{CNL} is the energy location of E_{CNL} from the vacuum level, E_{VAC} , S is the slope factor, ϵ_0 is the permittivity of vacuum, ϵ_i is the relative permittivity of the interface, q is the electronic charge,

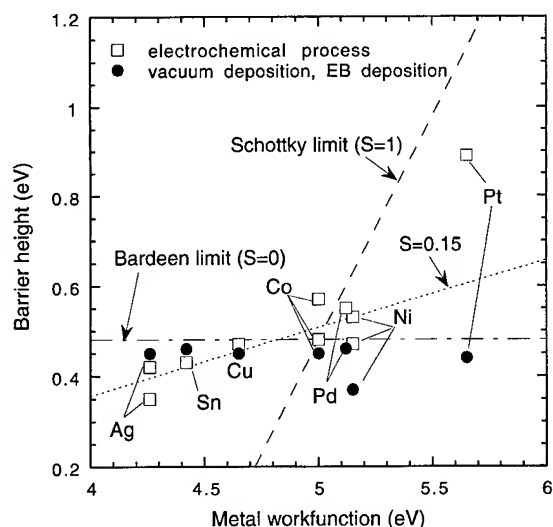


FIG. 10. Metal work function dependence of SBHs of metal/*n*-InP contacts formed by the *in situ* electrochemical process and vacuum evaporation process. The dotted line indicates the slope obtained from the least square fitting of the electrochemical data except for the Pt.

$D_{\text{MIGS}}(E_{\text{CNL}})$ is the density of MIGS at E_{CNL} , and δ_i is the thickness of the interface dipole layer (typically 0.2 nm) determined by the penetration depth of MIGS. Here, notations are changed from Ref. 18 so as to use work function rather than electronegativity. As for the actual value of S , Mönch has recently arrived at the following semiempirical formula:¹⁸

$$S \cong \frac{1}{1 + 0.29 \frac{(\epsilon_\infty - 1)^2}{\epsilon_i}} \quad (4)$$

where ϵ_∞ is the electronic permittivity of semiconductor. Using $\phi_m = 5.65$ eV,¹⁵ $\phi_{\text{CNL}} = 5.0$ eV,¹⁸ $\epsilon_\infty = 9.5$ ¹⁹ for InP, and $\epsilon_i = 4$,¹⁸ one obtains $S = 0.16$ and the value of the correction term of $S(\phi_m - \phi_{\text{CNL}})$ in Eq. (2) of 104 meV. However, this magnitude of correction is obviously too small to explain the experimental data.

Thus, the MIGS model which should be applicable to an ideal intimate contact can explain the SBH values of the various less ideal Schottky diodes with stress, lateral nonuniformity, and interfacial layer, but ironically not at all explains the SBH of the most ideal diode in this study which has an oxide-free, stress-free, defect-free, and intimate interface.

Then, from the viewpoint of the MIGS model, the following two possible explanations based on presence of an interfacial layer should be further investigated to understand why the MIGS model fails here. One possibility is that, although the intrinsic SBH is indeed determined by the penetration of MIGS, a certain interfacial dipole layer is formed by Pt electrodeposition which gives rise to additional dipole contribution to increase SBH. The other is, conversely, that an interfacial layer is formed so as to attenuate the evanescent metal wave function sufficiently before it forms MIGS in the semi-

conductor. In this case, SBH approaches the ideal Schottky limit, and should increase in the case of Pt due to its large work function. In either case, the interfacial layer is expected to be ultrathin, since the penetration depth of MIGS is extremely small, being usually below 1 nm.

One possible candidate for such an interfacial layer is some kind of oxide layer formed during the electrodeposition of Pt which provides an additional dipole to enhance the SBH, or attenuate MIGS. However, our XPS in-depth profile analysis of the present interface showed the absence of such an oxide layer down to the monolayer level, as already mentioned previously.

Another possible candidate for the interfacial layer is a hydrogen related layer, since electrodeposition is a competing process with hydrogen evolution. If such a layer exists, it may add additional dipole or screen MIGS. In fact, Wittmer and Freeouf²⁰ showed previously that an ideal Schottky diode whose SBH is far away from the prediction by the MIGS model can be realized by putting mercury contacts on hydrogen-passivated silicon surfaces. Since detection of hydrogen related ultrathin layer is extremely difficult in the present system, we cannot entirely rule out such a possibility. However the fact that the present Pt deposition process has a high current efficiency, and that the resultant electroplated Pt contacts adhere to InP very strongly and thermally stable at least up to 250–300 °C, seem to indicate that presence of such a weakly bonded layer which prevents direct interaction between metal and semiconductor, is highly unlikely.

Furthermore, excellent agreement of SBH values obtained between *I*-*V* and *C*-*V* methods as mentioned previously also indicates consistently that no interfacial layer exists at the interface. Additionally, $1/C^2$ -*V* plots were straight lines with no change of bulk doping near interface by the electrodeposition, indicating that no hydrogen passivation of dopants takes place. Thus, the present high SBH is an intrinsic property of an intimate Pt Schottky contact without the interfacial layer.

Thus in order to explain the present result, we have to investigate other models such as UDM, EWF model, and DIGS model all of which assert that Fermi level pinning is extrinsic, rather than intrinsic as in the case of the MIGS model, with different extrinsic origins for pinning.

From the viewpoint of UDM, the present result could be explained by reduction of the density of the discrete defect level responsible for pinning. However, in contrast to the case of GaAs where the well-known EL2 midgap level exists and is claimed to be the pinning center,²¹ no such defect which can cause pinning at 400–500 meV from E_c exists in InP, and this makes further discussion difficult.

From the viewpoint of EWF model, the present result may be explained if the electrodeposited Pt has an effective work function higher than that of the ordinary Pt. Since Pt is catalytically active and can dissolve appreciable quantities of hydrogen, its work function may be changed due to hydrogen evolution and subsequent incorporation during Pt electrodeposition. In fact, effects of hydrogen exposure on *n*-GaAs/Pt group metal Schottky barriers were previously

studied systematically by Aspnes and Heller,²² and it was demonstrated that hydrogen exposure could change SBH. However, the effect was to reduce SBH and not to increase. Thus, the EWF model does not seem to explain the present result either. Additionally, the reproducibility and thermal stability of the present Pt contacts also seem to indicate that such a possibility is unlikely.

We believe that the present result can be consistently explained by the disorder-induced gap state (DIGS) model for Fermi level pinning⁴⁻⁶ proposed by our group. According to this model, deposition of metal or insulator on the semiconductor surface disturbs the crystalline perfection of semiconductor surface and forms a thin disordered semiconductor layer with a thickness δ whose electronic properties are featured by the DIGS continuum. The continuum has continuous energy and spatial distribution of gap states of acceptor type and donor type with a characteristic charge neutrality level E_{HO} . E_{HO} agrees with that of the MIGS model, E_{CNL} , and lies at about 5.0 eV from the vacuum level for a major semiconductor.⁴

The DIGS model gives the same formula of Eq. (2) for the SBH, but, with the following expression for the slope factor.^{5,6}

$$S = \text{sech}(\delta/\lambda) \quad (5)$$

with

$$\lambda = [\epsilon_0 \epsilon_\infty / q^2 N_{\text{DIGS}}(E_{HO})]^{1/2}, \quad (6)$$

where λ is the DIGS screening length and $N_{\text{DIGS}}(E_{HO})$ is the volume density of DIGS at the charge neutrality level E_{HO} . According to this model, Fermi level pinning is extrinsic and the value of S is process dependent. When the degree of disorder is high, S becomes zero and the Fermi level is firmly pinned at the same position with the MIGS model. However, when an intimate M-S interface is formed without causing disorder, S becomes unity and the ideal Schottky limit is realized. These two limits are shown by dashed lines in Fig. 10, assuming $E_{HO} = E_V + 0.87 \text{ eV} = E_{\text{VAC}} - 5.0 \text{ eV}$. According to this, a large value of S of 0.62 was realized by electrodeposition of Pt.

In the present study, a high processing energy is used for formation of metal-semiconductor interface in the standard EB deposition and rf-sputtering deposition. Furthermore, random stress is generated at the interface due to large grain sizes and island formation in the case of electroplating of Co, Ni, and Ag. In these cases, a large disorder is produced at the semiconductor surface and produces high-density DIGS states which pin the Fermi level at E_{HO} , being independent of the metal work function. On the other hand, the present *in situ* electrochemical process possesses an extremely low processing energy with an application of small voltages of 500–700 mV at room temperature, and realizes for Pt deposition a nm size grain based stress-free intimate interface. This realizes a reduced DIGS density and realizes a large value of S .

V. CONCLUSION

In an attempt to understand the underlying mechanism of the recently observed large SBH enhancement by electrodeposition of Pt on InP, the evolution process of the metal (Pt, Ni, Co, and Ag)-InP interface was studied in detail, using I - V , C - V , DLTS, XPS, Raman, AFM, and SEM measurements.

It is shown that Pt deposition by the electrochemical process realizes an oxide-free, defect-free, stress-free, and nearly pinning-free ideal interface which has a remarkably high SBH of 900 mV.

On the other hand, Pt deposition by conventional vacuum deposition processes as well as Ag, Ni, and Co deposition by the electrochemical process lead to stressed and pinned interfaces with almost metal-independent SBH values in the range of 400–500 mV.

The observed result cannot be explained by the currently predominant metal induced gap state (MIGS) model, even if one takes account of the limited screening capability of the MIGS continuum. It can be explained neither by the unified defect model (UDM), nor by the effective work function (EWF) model.

It was shown that the results can be phenomenologically explained in a consistent way by the disorder induced gap state (DIGS) model proposed by the present author's group.

Further work is necessary, however, to directly verify the existence of DIGS, and establish the quantitative validity of the formula of Eq. (5). Since many theoretical calculation show existence of MIGS, a further work is also necessary to reconcile MIGS and DIGS models. Our present interpretation is that the MIGS will be definitely present, but its screening capability of the contact potential difference by an extremely thin dipole layer is somehow overestimated. Thus, the usually observed strong pinning is not due to MIGS, but due to DIGS. From this viewpoint, to what extent one can approach the Schottky limit is eventually determined by the screening capability of intrinsic MIGS. Further study may clarify this important question.

ACKNOWLEDGMENTS

The authors would like to thank Professor W. Mönch for fruitful discussion. The authors would also like to thank Dr. A. Noya and Dr. M. Takeyama of Kitami Institute Technology for fabrication of Pt/*n*-InP diodes by sputtering. This work was supported in part by a Grant-in Aid (Nos. 07455017, 07555093, and 08555072) from the Ministry of Education, Science, Culture, and Sports.

¹W. E. Spicer, I. Lindau, P. Skeath, and C. Y. Su, *J. Vac. Sci. Technol.* **16**, 1422 (1979).

²V. Heine, *Phys. Rev. Lett.* **A 138**, 1689 (1965).

³J. Tersoff, *Phys. Rev. Lett.* **52**, 465 (1984).

⁴H. Hasegawa and H. Ohno, *J. Vac. Sci. Technol. B* **4**, 1130 (1986).

⁵H. Hasegawa, *Proceedings of the 18th International Conference on the Physics of Semiconductors* (1986), Vol. 1, p. 291.

⁶H. Hasegawa, *Metal-Semiconductor Interfaces*, edited by A. Hiraki (IOS Tokyo, 1995), p. 280.

⁷J. M. Woodall and J. L. Freeouf, *J. Vac. Sci. Technol.* **19**, 794 (1981).

- ⁸N. J. Wu, T. Hashizume, H. Hasegawa, and Y. Amemiya, *Jpn. J. Appl. Phys.* **34**, 936 (1994).
- ⁹T. Sato, S. Uno, T. Hashizume, and H. Hasegawa, *Jpn. J. Appl. Phys.* **36**, 106 (1997).
- ¹⁰S. Uno, T. Hashizume, S. Kasai, N. J. Wu, and H. Hasegawa, *Jpn. J. Appl. Phys.* **35**, 1258 (1996).
- ¹¹T. Hashizume, H. Hasegawa, T. Sawada, A. Grub, and H. L. Hartnagel, *Jpn. J. Appl. Phys.* **32**, 486 (1993).
- ¹²T. Hashizume, G. Schweeger, N. J. Wu, and H. Hasegawa, *J. Vac. Sci. Technol. B* **12**, 2660 (1994).
- ¹³E. R. Rhoderick and R. H. Williams, *Metal-Semiconductor Contacts*, 2nd ed. (Oxford University Press, Oxford, 1978), p. 120.
- ¹⁴H. Norde, *J. Appl. Phys.* **50**, 5052 (1979).
- ¹⁵H. B. Michaelson, *J. Appl. Phys.* **48**, 4729 (1977).
- ¹⁶J. Tersoff, *J. Vac. Sci. Technol. B* **4**, 1066 (1986).
- ¹⁷M. Cardona and N. E. Christensen, *Phys. Rev. B* **35**, 6182 (1987).
- ¹⁸W. Mönch, *Appl. Surf. Sci.* **92**, 367 (1996).
- ¹⁹"Properties of Indium Phosphide", emis data reviews series no. 6, INSPEC, 1991, p. 110, an average of Table I was taken.
- ²⁰M. Wittmer and J. L. Freecouf, *Phys. Rev. Lett.* **69**, 2701 (1992).
- ²¹W. E. Spicer, R. Cao, K. Miyano, T. Kendelewicz, I. Lindau, E. Weber, Z. Liliental-Weber, and N. Newman, *Appl. Surf. Sci.* **41/42**, 1 (1989).
- ²²D. E. Aspnes and A. Heller, *J. Vac. Sci. Technol. B* **1**, 602 (1983).

Comparison of electron affinity and Schottky barrier height of zirconium and copper–diamond interfaces

P. K. Baumann^{a)} and R. J. Nemanich

Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202

(Received 5 February 1997; accepted 29 April 1997)

In this study, the evolution from diamond surfaces to metal–diamond interfaces has been examined. The electron affinity and the Schottky barrier height of a few Å thick films of Zr and Cu deposited in ultrahigh vacuum (UHV) onto IIb substrates were correlated. Prior to metal deposition, the diamond surfaces have been cleaned by different anneals and plasma treatments in UHV, and the surfaces were characterized by Auger electron spectroscopy and atomic force microscopy. The initial surfaces were terminated with oxygen, or free of chemisorbed species. Ultraviolet photoemission spectroscopy was employed to determine whether the samples exhibited a positive electron affinity or a negative electron affinity (NEA) before and after metal deposition. For Zr, the Schottky barrier height was found to change very little with the presence or absence of chemisorbed species at the interface. A NEA was observed for Zr on diamond independent of the surface termination. However, for Cu, the surface cleaning prior to metal deposition had a more significant effect. The Schottky barrier height changed strongly depending on the chemical species at the interface. A NEA was only detected for Cu on clean diamond surfaces. The differences between Zr on the one hand and Cu on the other are correlated with differences in interface chemistry and structure. © 1997 American Vacuum Society. [S0734-211X(97)11604-3]

I. INTRODUCTION

Negative electron affinity (NEA) surfaces could enable the development of cold cathode devices. The electron affinity of a semiconductor corresponds to the energy difference between the vacuum level and the conduction-band minimum. For most materials, the vacuum level lies above the conduction-band minimum. This is called a positive electron affinity. As a wide bandgap semiconductor, diamond has the potential of exhibiting negative electron affinity surfaces since the conduction-band minimum lies near the vacuum level. Electrons from the conduction-band minimum then have sufficient energy to leave a NEA surface and be emitted into vacuum.

By employing different surface treatments, such as plasma cleaning or annealing in ultrahigh vacuum (UHV), the position of the conduction-band minimum can be shifted with respect to the vacuum level. This can induce a NEA or remove it.^{1–6} Subsequent to precleaning the diamond (100) samples with a wet chemical etch, the diamond surfaces are oxygen terminated. This chemisorbed oxygen layer forms a surface dipole. Such a surface exhibits a positive electron affinity. For the diamond (100) surface, an anneal to 900–1050 °C or a H-plasma clean results in a NEA and a 2×1 reconstructed, oxygen-free surface.^{3,5,6} The different threshold temperatures are related to different wet chemical pretreatments.³ It was found that UHV annealing at 900 °C was sufficient for samples precleaned by an electrochemical etch. But a 1050 °C anneal was required for a preclean employing chromic acid. It has been proposed that the diamond (100) surface exhibits a monohydride termination subsequent to a 900–1050 °C anneal or a H-plasma exposure.^{3,5,6} A H

surface layer results in a dipole resulting in a NEA. However, for all these treatments, a positive electron affinity and a 2×1 reconstruction are observed following a 1150 °C anneal.⁶ This surface is considered to be free of adsorbates. *Ab initio* calculations for the 2×1 reconstructed surface predict a NEA for a monohydride terminated surface and a positive electron affinity for an adsorbate-free surface.⁵ This is in agreement with the experimental results.^{5,6}

Various studies have described metals on diamond.^{7–15} As-deposited metal contacts usually have been found to exhibit Schottky characteristics. Most of these reports are based on current–voltage (*I*–*V*) data. It is, however, difficult to deduce the Schottky barrier height from *I*–*V* measurements due to the large ideality factors of metal–diamond interfaces.^{13–15} Ultraviolet photoemission spectra can be used to measure the Schottky barrier height of metal–diamond interfaces. For this purpose, features from both the metal and the semiconductor need to be visible. This means the thickness of the metal has to be equal to or less than the electron mean-free path (≤ 5 Å). It has been demonstrated that depositing a few Å of metals such as Ti, Ni, Co, Cu, and Zr can induce a NEA on diamond surfaces.^{3,16–21} Also, sub-monolayer deposition of TiO has been reported to induce a NEA.²² The presence of a NEA or positive electron affinity has been correlated with different structures of the metal–diamond interface. Metal films deposited on adsorbate-free surfaces have been found to exhibit lower Schottky barrier heights and lower electron affinities than for surfaces terminated by species such as hydrogen or oxygen. And for some metal–diamond structures, the Schottky barrier heights have been low enough to induce a NEA.

Photoemission spectroscopy is a very sensitive technique to determine whether a surface exhibits a NEA or a positive

^{a)}Electronic mail: robert_nemanich@ncsu.edu

electron affinity. Electrons are photoexcited from the valence band into states in the conduction band. These electrons can then quasithermalize down to the conduction-band minimum. For NEA surfaces, these secondary electrons from the conduction-band minimum can be emitted into vacuum and appear as a sharp peak at the low kinetic energy end of the photoemission spectra.^{23,24} In this study, diamond (100) surfaces have been cleaned by anneals to 1050 or 500 °C. Thin Zr or Cu films were deposited on these diamond substrates. The surface properties were analyzed before and after metal deposition.

II. EXPERIMENT

For this study, an UHV system was employed that consists of several interconnected chambers featuring capabilities for annealing, metal deposition, ultraviolet photoemission spectroscopy (UPS), and Auger electron spectroscopy (AES). Several natural type IIb single-crystal semiconducting, boron doped diamond (100) substrates (3.0×3.0×0.25 mm) were used. Typical resistivities of these samples were $10^4 \Omega \text{ cm}$. To remove nondiamond carbon and metal contaminants, an electrochemical etch has been employed. Details of this method have been previously described.²⁵ The wafers were blown dry with N_2 , mounted on a Mo holder and transferred into the UHV system (base pressure $\sim 1 \times 10^{-10}$ Torr). Two different *in vacuo* cleaning processes were employed to study the effect of surface treatment on the characteristics of the metal–diamond interface. These processes consisted of an anneal to either 1150 or 500 °C both for 10 min. The pressure in the annealing chamber rose from 1×10^{-10} to 7×10^{-9} and 8×10^{-10} Torr during the anneals, respectively. After the heat treatment, 2 Å thick layers of Zr or Cu were deposited onto the diamond surface. The deposition was facilitated by an e-beam evaporator. And the diamond substrates were kept at room temperature during metal deposition. The pressure in the chamber rose to 2×10^{-9} Torr during deposition. A quartz-crystal monitor was employed to determine the thickness of the metal films. Following the annealing, and the growth steps, UPS and AES were employed to characterize the surface properties.

The presence of Zr or Cu on the surface was confirmed by using AES. Atomic force microscopy (AFM) images of the diamond wafers clearly showed arrays of linear grooves parallel to each other with a depth of ~ 20 Å. This surface structure is a result of the commercial polishing procedure used to smoothen the surfaces. Subsequent to depositing 2 Å of metal, no island structures were observed in the AFM measurements. This is indicative of a uniform two-dimensional layer for both Zr and Cu.

He I (21.21 eV) radiation from a discharge lamp was employed to facilitate the photoemission. The emitted electrons were measured using a 50 mm hemispherical analyzer with an energy resolution of 0.15 eV. The sample was biased by 1 V with respect to the analyzer. This was necessary so that the low-energy electrons from the NEA surfaces could be detected, despite the work function of the analyzer. The sharp

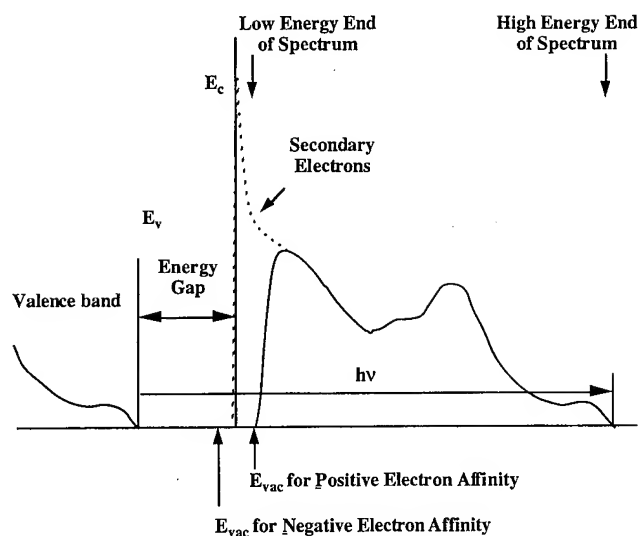


FIG. 1. Schematic diagram of photoemission spectra for a negative electron affinity surface (dotted line) and a positive electron affinity surface (solid line).

NEA peak appears at the low-energy end of the photoemission spectrum and corresponds to the energy position of the conduction-band minimum E_C (Fig. 1). Emission from E_C is positioned at $E_V + E_G$ in the spectrum. E_V is the energy position of the valence-band maximum and E_G the energy of the band gap. In a corresponding manner, emission from E_V appears at $E_V + h\nu$ in the UPS spectra. For semiconductors, this corresponds to the high kinetic energy cutoff of the spectra. Then, the spectral width for NEA surfaces, or the distance between emission from the valence-band maximum and the conduction-band minimum is given by $h\nu - E_G$. With the values for He I radiation $h\nu = 21.21$ eV and the band gap of diamond $E_G = 5.47$ eV, the width of the spectrum is 15.7 eV. However, for a surface with a positive electron affinity, the low-energy cutoff is determined by the vacuum level. Then, the spectral width will be smaller than for a NEA. In fact, the width will be reduced by the value of the positive electron affinity of the surface.

Consider photoemission spectra of a thin metal film on a semiconductor. Spectra exhibiting features from both the metal and the semiconductor can be used to determine the Schottky barrier height Φ_B (Fig. 2). For this to be the case, the thickness of the metal layer needs to be equal to or less than the electron mean-free path (≤ 5 Å). The Schottky barrier height for a metal on a *p*-type semiconductor, like diamond, is defined by the difference between the position of the Fermi level of the metal, E_F and the valence-band edge, E_V , of the semiconductor. It may, however, be difficult to detect the relatively weak onset of emission at E_V , even for metal layers thinner than the mean-free path. This is due to the fact that emission from the metal Fermi level may overshadow this weak onset. Often times it is, therefore, necessary to use an independent method to determine the position of the valence-band minimum. In particular, E_V can be referenced to some strong features in the diamond spectrum before metal deposition. Here, a peak positioned at 8.3 eV

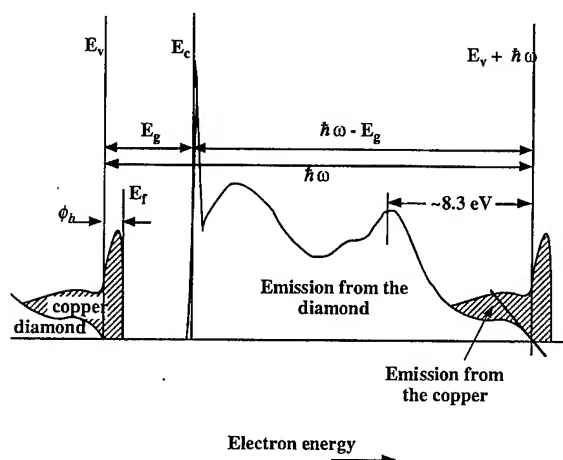


FIG. 2. Schematic diagram of photoemission spectra for a metal (e.g., copper) deposited on diamond. The Schottky barrier height Φ_B is determined from the difference between the position of the valence-band edge of diamond E_V and the metal Fermi-level E_F .

below E_V was chosen as a reference. In the case of a NEA surface, the sharp low-energy peak corresponding to E_C can be used to find E_V , too. Then the difference between E_C and E_V corresponds to $h\nu - E_G$.

III. RESULTS AND DISCUSSION

First consider the diamond surfaces before metal deposition. After loading the samples into the vacuum system, AES scans were obtained. Features indicative of the presence of oxygen on the surface were detected. Subsequent to a 500 °C anneal, the oxygen peak was only slightly reduced. After a 1150 °C anneal, oxygen was no longer observed by means of AES. From UPS spectra, positive electron affinities of $\chi=1.4$ eV and of $\chi=0.7$ eV were found for the substrates heated to 500 and 1150 °C, respectively. These numbers are consistent

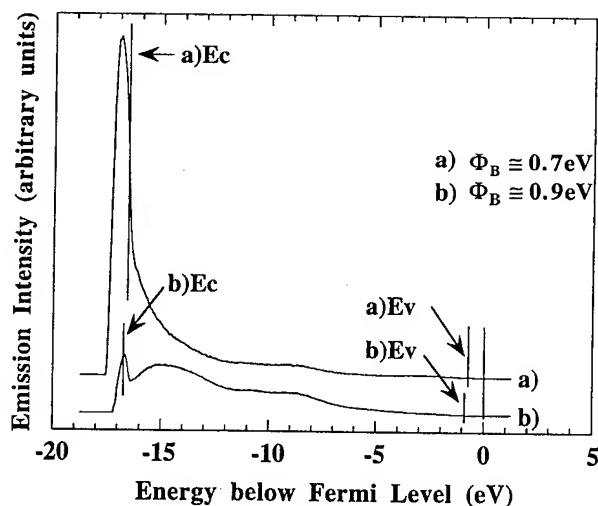


FIG. 3. UV photoemission spectra of 2 Å of Zr deposited on a diamond (100) surface annealed to (a) 1150 °C and (b) 500 °C. Metal induced NEAs are observed upon deposition of Zr for both (a) and (b). For (a), emission below E_C is detected.

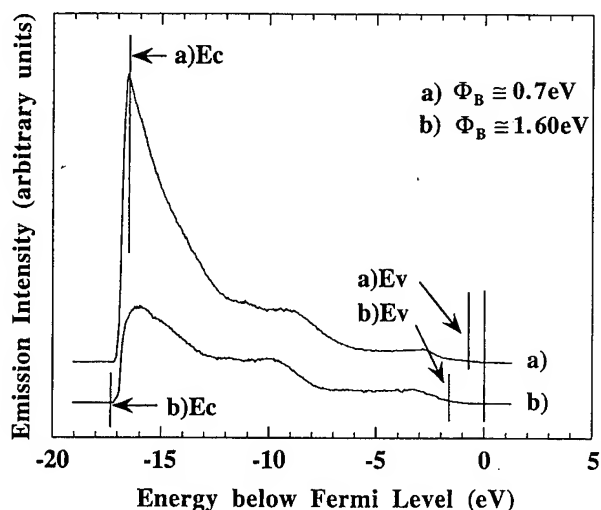


FIG. 4. UV photoemission spectra of 2 Å of Cu grown on a diamond (100) surface following an anneal to (a) 1150 °C and (b) 500 °C. A metal induced NEA is observed for (a), whereas a positive electron affinity is detected for (b).

with values reported previously.^{5,18} It is expected that chemisorbed oxygen on diamond results in a stronger surface dipole than for the clean surface. This would also lead to a larger work function for the oxygen terminated surface. Our observations are in agreement with that.

Following deposition of 2 Å of Zr on clean diamond surfaces, the width of the UPS spectrum increased consistently with the surface, exhibiting a NEA [Fig. 3(a)]. A Schottky barrier height of $\Phi_B=0.70$ eV was measured. Emission below the conduction-band minimum E_C was observed. This phenomenon will be discussed further in another publication.²⁶ Deposition of 2 Å of Zr on oxygen terminated diamond surfaces also resulted in a NEA. A larger Schottky barrier height of $\Phi_B \approx 0.9$ eV was measured [Fig. 3(b)]. And the spectrum shifted ~ 0.3 eV toward lower energies. Subsequent to depositing 2 Å of Cu on clean diamond surfaces, a NEA and a Schottky barrier height of $\Phi_B=0.70$ eV were determined by means of UPS [Fig. 4(a)]. Also, the spectra shifted by 0.3 eV to lower energies. However, in the case of Cu on oxygen terminated surfaces, a positive electron affinity of $\chi=0.75$ eV and a larger Schottky barrier height of $\Phi_B \approx 1.60$ eV were measured [Fig. 4(b)]. Also, a larger shift to lower energies of 0.6 eV was found. These results are summarized in Table I. Equation (1) is valid specifically for photoemission of thin metal layers (less than the electron

TABLE I. Results of UV photoelectron emission spectroscopy to measure electron affinity χ and Schottky barrier Φ_B . PEA: positive electron affinity, NEA: negative electron affinity.

Sample	UPS oxygen	Clean
C(100)	PEA, $\chi \approx 1.40$ eV	PEA, $\chi \approx 0.70$ eV
Zr/C(100)	NEA, $\chi < 0$, $\Phi_B \approx 0.90$ eV	NEA, $\chi < 0$, $\Phi_B \approx 0.70$ eV
Cu/C(100)	PEA, $\chi \approx 0.75$ eV, $\Phi_B \approx 1.60$ eV	NEA, $\chi < 0$, $\Phi_B \approx 0.70$ eV

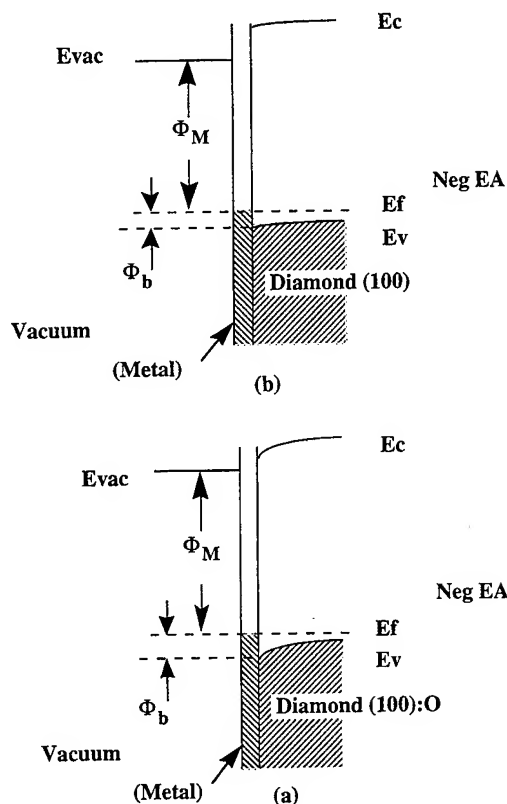


FIG. 5. Band diagrams of the zirconium-diamond interface. For zirconium, on both the oxygen terminated (a) and the clean surface (b), the Schottky barrier height added to the metal work function is less than the diamond band gap. This corresponds to a NEA.

mean-free path) on semiconductors: The electron affinity can be expressed in terms of the Schottky barrier formed with a *p*-type semiconductor.²⁷

$$\chi = (\Phi_M + \Phi_B) - E_G. \quad (1)$$

Using the band gap of diamond $E_G = 5.47$ eV, the work function of Zr ($\Phi_M = 4.05$ eV) and Cu ($\Phi_M = 4.59$ eV), and the measured Schottky barrier heights Φ_B , the electron affinities can be calculated. For Zr, we obtain $\chi = -0.72$ eV for the clean surface and $\chi = -0.52$ eV for the oxygen terminated surface. In the same way, $\chi = -0.18$ eV and $\chi = 0.72$ eV are obtained for Cu on the clean and oxygenated surfaces, respectively. These results are consistent with the NEA and positive electron affinity effects that were observed by employing the UPS. Figures 5 and 6 show energy-band diagrams of the Zr-diamond and the Cu-diamond interfaces. These schematics illustrate the correlation of the Schottky barrier height with the electron affinity.

This simple work-function model has been used successfully to explain NEA or positive electron affinity effects for systems like Ti or Ni layers on diamond (111) surfaces.^{16,17} Other approaches may be used to interpret these results. The monolayer thick metal films could be considered as dipole layers. It has been found that Ni deposited on Ar-plasma cleaned diamond (111) substrates induces a NEA. An Ar plasma or a 950 °C anneal results in a (111) surface free of adsorbates.⁴ In comparison, a larger Schottky barrier and a

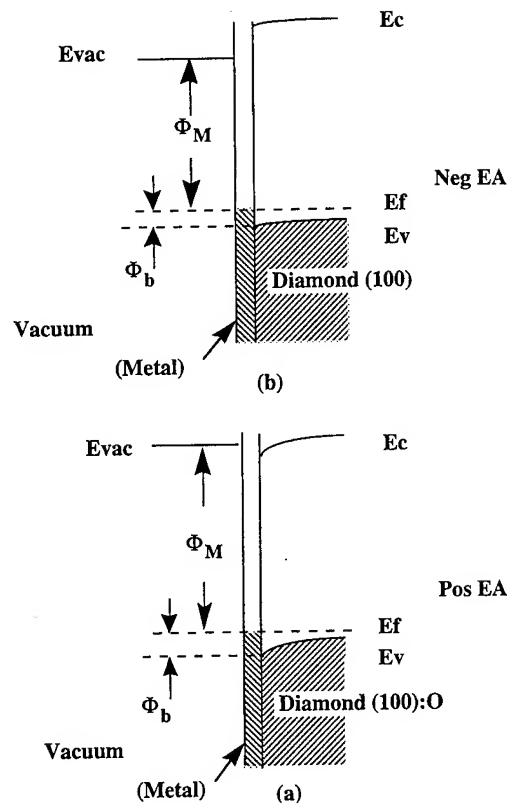


FIG. 6. Band diagrams of the copper-diamond interface. For copper on the oxygen terminated surface (a), the sum of the Schottky barrier height and work function for metal on diamond is greater than the band gap of diamond, resulting in a positive electron affinity. For copper on the clean surface (b), the Schottky barrier height added to the metal work function is less than the diamond band gap. This corresponds to a NEA.

positive electron affinity were measured for thin Ni films on hydrogen terminated (111) surfaces. In theoretical studies by Erwin and Pickett²⁸⁻³¹ and Pickett, Pederson, and Erwin,³² it was reported that the most stable configuration for Ni on clean (111) and (100) surfaces resulted in a Schottky barrier height of less than 0.1 eV. Considering copper on diamond (111) surfaces, Lambrecht³³ calculated a value for the Schottky barrier height of less than 0.1 eV for clean surfaces and greater than 1.0 eV for hydrogen terminated surfaces. According to these results, the interface termination appears to have a significant effect on the Schottky barrier height. For metals deposited on clean surfaces, lower values for the Schottky barrier height and a greater likelihood of inducing a NEA are expected than for metals on non-adsorbate-free surfaces. The Schottky barrier heights reported in our study for Zr and Cu on diamond are consistent with this.

Apparently, the Schottky barrier height for Zr on diamond does not depend on the surface termination of the diamond substrate as strongly as is the case for Cu. Both Zr on clean, as well as oxygen terminated diamond surfaces, tend to exhibit lower electron affinities than Cu on corresponding surfaces. This could be due to the higher reactivity of Zr with both C and O than Cu. It has been reported that Ti, as well as titanium oxide on diamond, exhibit a NEA.²² Zr is next to Ti in the periodic table of elements and has properties similar to

Ti. In our experiments, Zr could have reacted with the oxygen from the oxygen terminated surface. This may be indicative that Zr, as well as zirconium oxide on diamond, could exhibit a NEA. Zr, like Ti, does react with C. But this reaction is not expected to occur at room temperature. Ti was annealed to $>400^\circ\text{C}$ before reaction with C was observed.³⁴ Cu, on the other hand, does not react as readily with C or O. Thus, the Cu–diamond interface structure for Cu on the clean diamond surfaces is different than for Cu on the oxygen terminated surfaces.

IV. CONCLUSIONS

The effects of depositing thin metal films onto clean and oxygen terminated diamond (100) substrates has been studied by UPS. It was found that Cu induced a NEA on clean surfaces but not on oxygen terminated surfaces. In comparison, Zr induced a NEA on both clean and oxygen terminated surfaces. The Schottky barrier height of Zr on diamond was less dependent on the termination of the diamond surface than was the case for Cu. This is attributed to the fact that Zr exhibits a strong affinity to the oxygen of the oxygen terminated diamond surfaces. In comparison, Cu does not exhibit a significant tendency to form oxides.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research (Contract No. N00014-92-J-1477).

¹F. J. Himpsel, D. E. Eastman, P. Heimann, and J. F. van der Veen, *Phys. Rev. B* **24**, 7270 (1981).

²B. B. Pate, M. H. Hecht, C. Binns, I. Lindau, and W. E. Spicer, *J. Vac. Sci. Technol.* **21**, 364 (1982).

³P. K. Baumann, T. P. Humphreys, and R. J. Nemanich, *Mater. Res. Soc. Symp. Proc.* **339**, 69 (1994).

⁴J. van der Weide and R. J. Nemanich, *Appl. Phys. Lett.* **62**, 1878 (1993).

⁵J. van der Weide, Z. Zhang, P. K. Baumann, M. G. Wensell, J. Bernholc, and R. J. Nemanich, *Phys. Rev. B* **50**, 5803 (1994).

⁶P. K. Baumann and R. J. Nemanich, *Proceedings of Diamond Films '94 of the 5th European Conference on Diamond, Diamond-like and Related Materials*, edited by P. K. Bachmann, I. M. Buckley-Golder, J. T. Glass, and M. Kamo, *J. Diam. Relat. Mater.* **4**, 802 (1995).

⁷K. Das, V. Venkatesan, K. Miyata, D. L. Dreifus, and J. T. Glass, *Thin Solid Films* **212**, 19 (1992).

⁸Y. Mori, H. Kawanada, and A. Hiraki, *Appl. Phys. Lett.* **58**, 940 (1991).

⁹J. F. Prins, *J. Phys. D* **22**, 1562 (1989).

¹⁰H. Shiomi, H. Nakahata, T. Imai, Y. Nishibayashi, and N. Fujimori, *Jpn. J. Appl. Phys.* **28**, 758 (1989).

¹¹T. Tachibana, B. E. Williams, and J. T. Glass, *Phys. Rev. B* **45**, 11 975 (1992).

¹²M. W. Geis, D. D. Rathman, D. J. Ehrlich, R. A. Murphy, and W. T. Lindley, *IEEE Electron Device Lett.* **8**, 341 (1987).

¹³M. C. Hicks, C. R. Wronski, S. A. Grot, G. S. Gildenblat, A. R. Badzian, T. Badzian, and R. Messier, *J. Appl. Phys.* **65**, 2139 (1989).

¹⁴G. Zhao, T. Stacy, E. J. Charlson, E. M. Charlson, C. H. Chao, M. Haj-said, and J. Meese, *Appl. Phys. Lett.* **61**, 1119 (1992).

¹⁵D. G. Jeng, H. S. Tuan, R. F. Salat, and G. J. Fricano, *J. Appl. Phys.* **68**, 5902 (1990).

¹⁶J. van der Weide and R. J. Nemanich, *J. Vac. Sci. Technol. B* **10**, 1940 (1992).

¹⁷J. van der Weide and R. J. Nemanich, *Phys. Rev. B* **49**, 13 629 (1994).

¹⁸P. K. Baumann and R. J. Nemanich, *Appl. Surf. Sci.* **104/105**, 267 (1996).

¹⁹P. K. Baumann and R. J. Nemanich, *Mater. Res. Soc. Symp. Proc.* **416**, 157 (1996).

²⁰P. K. Baumann, S. P. Bozeman, B. L. Ward, and R. J. Nemanich, *Mater. Res. Soc. Symp. Proc.* **423**, 143 (1996).

²¹P. K. Baumann, S. P. Bozeman, B. L. Ward, and R. J. Nemanich, *Proceedings of Diamond '96, the 7th European Conference on Diamond, Diamond-like and Related Materials jointly with ICNDST-5, the 5th International Conference on the New Diamond Science and Technology*, edited by J. C. Angus, P. K. Bachmann, I. M. Buckley-Golder, O. Fukunaga, J. T. Glass, and M. Kamo, *J. Diam. Relat. Mater.* **6**, 398 (1997).

²²C. Bandis, D. Haggerty, and B. B. Pate, *Mater. Res. Soc. Symp. Proc.* **339**, 75 (1994).

²³F. J. Himpsel, P. Heimann, and D. E. Eastman, *Solid State Commun.* **36**, 631 (1980).

²⁴B. B. Pate, W. E. Spicer, T. Ohta, and I. Lindau, *J. Vac. Sci. Technol.* **17**, 1087 (1980).

²⁵M. Marchywka, P. E. Pehrsson, S. C. Binari, and D. Moses, *J. Electrochem. Soc.* **140**, L19 (1993).

²⁶P. K. Baumann and R. J. Nemanich (unpublished).

²⁷E. H. Rhoderick and R. H. Williams, *Metal-Semiconductor Contacts* (Clarendon, Oxford, 1988).

²⁸S. C. Erwin and W. E. Pickett, *Surf. Coat. Technol.* **47**, 487 (1991).

²⁹S. C. Erwin and W. E. Pickett, *Solid State Commun.* **81**, 891 (1992).

³⁰W. E. Pickett and S. C. Erwin, *Phys. Rev. B* **41**, 9756 (1990).

³¹W. E. Pickett and S. C. Erwin, *Superlattices Microstruct.* **7**, 335 (1990).

³²W. E. Pickett, M. R. Pederson, and S. C. Erwin, *Mater. Sci. Eng. B* **14**, 87 (1992).

³³W. R. L. Lambrecht, *Physica B* **185**, 512 (1993).

³⁴J. van der Weide and R. J. Nemanich, *Proceedings of the First International Conference on the Applications of Diamond Films and Related Materials*, edited by Y. Tzeng, M. Yoshikawa, M. Murakawa, and A. Feldman (Elsevier, New York, 1991), p. 359.

Nature and origins of stacking faults from a ZnSe/GaAs interface*

L. H. Kuo,^{a)} K. Kimura, A. Ohtake, and S. Miwa

*Joint Research Center for Atom Technology
and Angstrom Technology Partnership, 1-1-4 Higashi, Tsukuba 305, Japan*

T. Yasuda

*Joint Research Center for Atom Technology
and National Institute for Advanced Interdisciplinary Research, 1-1-4 Higashi, Tsukuba 305, Japan*

T. Yao

*Joint Research Center for Atom Technology
and National Institute for Advanced Interdisciplinary Research, 1-1-4 Higashi, Tsukuba 305, Japan
and Institute for Materials Research, Tohoku University, 2-1-1, Katahira, Aoba-ku, Sendai 980, Japan*

(Received 12 January 1997; accepted 25 March 1997)

Existence of ~ 3 – 4 monolayers of Ga_2Se_3 - and Ga_2Te_3 -like interfacial layers are suggested by transmission electron microscopy of Se- and Te-exposed (or -reacted) ZnSe/GaAs interfaces, respectively. Densities of extrinsic Shockley- and intrinsic Frank-type stacking faults are of $\sim 5 \times 10^7/\text{cm}^2$ in samples grown on Se- or Te-exposed GaAs surfaces. Annealing on the Se- or Te-exposed GaAs generated a high density of vacancy loops ($\geq 1 \times 10^9/\text{cm}^2$) with an increase of the densities of both intrinsic and extrinsic-type stacking faults ($\geq 5 \times 10^8/\text{cm}^2$) after growth of the films. Formation of the intrinsic stacking faults or vacancy loops and extrinsic stacking faults may be related to the presence of cation vacancies and interstitials, respectively, on the surface of the GaAs epilayer, due to the interaction between Se or Te and the GaAs epilayer with charge unbalanced Ga–Se or Ga–Te bondings. On the other hand, ~ 2 and 3 – 4 monolayers of Zn–As interfacial layers are recognized in samples grown on Zn-exposed GaAs-(2×4) and -c(4×4), respectively. A very low density of fault defects in the range of $\sim 10^4/\text{cm}^2$ was obtained in samples with Zn treatment on an As-stabilized GaAs-(2×4). However, the density of As precipitates increases as the surface coverage of c(4×4) reconstruction increased on the Zn-exposed As-stabilized GaAs-(2×4) surface and this is associated with an increase of the density of extrinsic-type stacking faults bound by partial edge dislocations with a core structure terminated on additional cations. Clustering of excess As atoms and formation of Zn interstitials due to charge unbalance of Zn–As bondings on the GaAs surface may act as nucleation sites for the generation of the high densities of As precipitates and extrinsic-type stacking faults, respectively. These local stacking errors due to the interaction between Zn and GaAs-c(4×4) is reduced by thermal annealing. In this case, thickness of the Zn–As interfacial layer is decreased with increasing annealing temperature and a very low density of the fault defects is obtained. © 1997 American Vacuum Society. [S0734-211X(97)09304-9]

I. INTRODUCTION

The chemistry of heterointerface plays an important role on the generation of crystalline defects especially in near lattice matched semiconductor heterostructure. In GaAlAs/GaAs,¹ InGaAsP/InP,² and ZnSe-based/GaAs heterostructures,^{3–8} the grown-in crystalline defects such as stacking faults, precipitates, vacancy loops, and misfit dislocations are sources for the generation of degradation defects from the active region that make the device fail during operation. In relaxed semiconductor heterostructures such as GeSi/Si,⁹ ZnSe/GaAs,^{10,11} and ZnTe/GaSb,¹² the fault defects including stacking faults, precipitates, and vacancy loops form below the critical thickness and act as low energy nucleation sites for the generation of misfit dislocations. In addition, these fault defects are blocking centers to prevent

gliding of the threading arms of the misfit dislocations.^{13–15} Consequently, a high density of pre-existing fault defects in the heterostructures gives rise to a high density of threading dislocations and short segments of misfit dislocations after the strain relaxation. Therefore, it is important to understand the origins and control the formation of these pre-existing fault defects to improve the performance of the semiconductor device.

From previous studies in ZnSe-based/GaAs heterostructure by molecular beam epitaxy (MBE)^{16–19} and organometallic vapor phase epitaxy,²⁰ the factors of interface chemistry on the generation of the grown-in fault defects include impurities, stoichiometries, and treatment on the surfaces of GaAs epilayers. In this fashion, growth of a GaAs epilayer on the GaAs substrate is essential to obtain a uniform interfacial stoichiometry and reduce the interfacial impurities or roughness. On the other hand, local incorporation between Zn and Se atoms under varied beam flux ratios (or growing surface stoichiometries) also plays an important role in the

*No proof corrections received from author prior to publication.

^{a)}Electronic mail: kuo@jrcat.or.jp

generation of fault defects in ZnSe/GaAs.^{19,21} In this case, generation of the stacking faults is attributed to the improper incorporation of growing species with a higher beam flux at the initial stages of growth.²¹

Structurally, formation of the stacking fault in zinc-blende structure is in the same as it is in the face-center-cubic (fcc) lattice. The zinc-blende structure can be described as an fcc lattice with a basis of two atoms per lattice point. Therefore, extrinsic fault and intrinsic fault involves insertion and removal of a pair of {111} planes (cation and anion lattice planes) to preserve tetrahedral coordination.²² In addition, since compound semiconductors have a zinc-blende structure consisting of two interpenetrating fcc sublattices (noncentrosymmetric arrangement), there are two types of partial edge dislocations, depending on where the extra half plane of the stacking faults lies.²³ For instance, stacking faults lying on (111) and ($\bar{1}\bar{1}\bar{1}$) planes are bound by partial edge dislocations with dislocation core structures terminated on cations and anions, respectively.

In this article we show a systematic study on the chemical composition of ZnSe/GaAs interface and the distribution of the stacking faults in the ZnSe films grown on GaAs epilayers as a function of chemical (Zn, Se, and Te) and thermal preparation on the surfaces of the GaAs epilayers. Different chemical composition of the ZnSe/GaAs interfaces and a strong dependence of the type and density of the stacking faults on the surface preparation of the GaAs epilayers is recognized.

II. EXPERIMENT

Thin pseudomorphic ZnSe films (~50–130 nm) were grown on GaAs epilayers on GaAs (001) substrates at 250–260 °C by MBE with II–VI and III–V MBE chambers connected by transfer modules under ultrahigh vacuum. The GaAs epilayers were grown at 520–540 °C with a growth rate of ~0.4–0.5 $\mu\text{m/h}$ and $P_{\text{As}}/P_{\text{Ga}}$ flux ratio of ~40. All the GaAs epilayers were grown under the same growth condition. The GaAs epilayers were grown and prepared by post-grown thermal treatments to obtain (2 \times 4), c(4 \times 4), a mixture of (2 \times 4) and c(4 \times 4), and (4 \times 6) surface reconstructions as evidenced by the observation of reflection high energy electron diffraction (RHEED) patterns and reflectance difference spectroscopy (RDS).²⁴ For the study of the defect generation related to the chemical preparation on the surfaces of the GaAs epilayers, the ZnSe films were grown at the same beam flux ratios (or growing surface stoichiometry) with a mixture of (2 \times 1) and weak c(2 \times 2) surface reconstruction. The surfaces of the As-rich GaAs-(2 \times 4) and -c(4 \times 4) epilayers are known to exist as single and double As-terminated layers, respectively.²⁵ In addition, the (4 \times 6) surface reconstruction is recognized as a Ga-rich GaAs surface.²⁵ Three sets of samples were grown with different chemical preparations of the GaAs surfaces. One set of samples was grown to study the effect of Se contamination or interaction with the GaAs substrates on the defect generation. Samples were grown with a 5 s Se exposure at 250–260 °C of the As-stabilized GaAs-(2 \times 4) surface and a 1 min

Se exposure of the Ga-rich GaAs-(4 \times 6) surface of the GaAs epilayers. In addition, thermal annealing at 520 °C was applied to the Se-exposed Ga-rich GaAs-(4 \times 6) to increase the interaction between the Se atoms and the GaAs surface prior to the growth of the thin ZnSe layer. In the case of the Se-treated or -reacted samples, a (2 \times 1) Se-rich GaAs surface reconstruction was observed by RHEED pattern prior to the growth of the ZnSe films. In addition, spotty RHEED reconstructions were observed at the initial stages of ZnSe growth on this Se-exposed or -reacted GaAs surface. This indicates that the island growth mode of ZnSe films was taking place on the (2 \times 1) Se-exposed GaAs surface.

Another set of samples was grown to study the effect of Te exposure and thermal annealing on the generation of the grow-in fault defects. In these samples, surfaces of As-stabilized GaAs-(2 \times 4) were exposed to Te at 300, 350, and 400 °C for 1–2 min and cooled down to growth temperature at ~260 °C prior to growth of the ZnSe films. For samples with a higher annealing temperature of ~500 °C, Te treatments were done at 400 °C on GaAs-(2 \times 4) for 1 min and annealed at 500 °C for 1 min. The surface reconstruction of the As-stabilized GaAs-(2 \times 4) was changed to GaAs-(6 \times 1) after the exposure of Te flux. In addition, one sample was grown on a GaAs surface with 1 min Zn and then Te treatments at 260 °C for comparison. In this case, the Zn treatments were applied on the GaAs-(2 \times 4) surface prior to the Te exposure to prevent the interaction between Te atoms and the GaAs-(2 \times 4) surface. Very streaky RHEED intensity oscillations were observed during the growth of the ZnSe films on the Te-exposed or -reacted GaAs surfaces.²⁶ This indicates that a layer-by-layer growth of the ZnSe films was achieved on the Te-exposed GaAs surfaces.

A third set of samples had 1 min Zn exposure at ~250–260 °C of (2 \times 4), a mixture of (2 \times 4) and c(4 \times 4), and c(4 \times 4) As-rich GaAs surfaces in order to study the effect of surface As coverage (or stoichiometry) on the generation of the grown-in fault defects. After 1 min of Zn exposure on an As-stabilized GaAs-(2 \times 4) surface, a half order of streaks in [110] azimuth of RHEED pattern became weaker. On the other hand, the c(4 \times 4) surface reconstruction disappeared and changed to (2 \times 2)-like after 1 min of Zn exposure at 260 °C. These results suggest that Zn indeed reacted or adhered with the As-stabilized GaAs-(2 \times 4) and c(4 \times 4) surfaces. In addition, thermal annealing at 400 and 440 °C under Zn exposure was applied on the Zn-exposed As-rich GaAs-c(4 \times 4) prior to the growth of the films. During the thermal annealing under Zn exposure, the surface reconstruction of Zn-exposed As-rich GaAs-c(4 \times 4) was changed to (2 \times 3)-like at ~320 °C and (2 \times 4)-like at ~420 °C by RHEED observation. The (2 \times 3) surface reconstruction remained unchanged within 10 min thermal annealing under Zn exposure at 390–400 °C. However, the c(4 \times 4) reconstruction is changed to (2 \times 4) within 2–3 min thermal annealing at ~400 °C without Zn exposure. Therefore, Zn exposure has an effect of hindering As desorption from the GaAs surface. In addition, the (2 \times 4)-like surface reconstruction

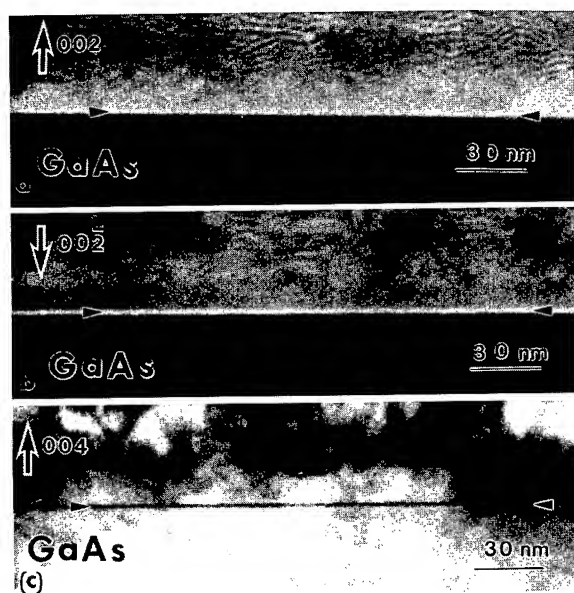


FIG. 1. TEM cross-sectional (a) 002, (b) $00\bar{2}$, and (c) 004 dark field images from a sample with 1 min Se exposure of a Ga-rich GaAs-(4 \times 6) epilayer prior to the growth of the ZnSe film. The Ga₂Se₃-like interfacial layer is marked by arrowheads.

tion remained the same during 5–70 min annealing at 440 °C under Zn exposure.

For transmission electron microscopy (TEM) studies, (110) and ($\bar{1}\bar{1}0$) cross-section specimens were prepared by mechanically thinning the samples using a tripod polisher followed by argon ion-milling at liquid nitrogen temperature. (001) plan-view specimens were prepared by etching the back side of the substrate using a 4:1 (H₂O₂:NH₄OH) solution. A JEOL 2010 TEM operated at 200 kV was used to study the structure and density of the defects. The density of stacking faults was carefully checked at many areas of the TEM specimens with well tilted 220-type reflection vectors. In addition, many well prepared TEM plan-view specimens were carefully checked to confirm the value of the low defect density. Specific directions in the ZnSe/GaAs are recognized by the etch-pit features on the back side of the GaAs substrates.

III. RESULTS AND DISCUSSION

A. Chemical composition of ZnSe/GaAs interface

1. Se-exposed or -reacted ZnSe/GaAs interfaces

TEM 002- and 004-type reflections were used to study chemical composition of the ZnSe/GaAs interfacial layers. We have studied the chemical composition of the interfacial layers with ZnSe films grown on the Se-exposed Ga-rich GaAs-(4 \times 6) surfaces with and without thermal annealing. In both cases, contrasts and thicknesses of the interfacial layers are roughly identical under 002- and 004-type reflections of the TEM dark field images as shown in Figs. 1(a), 1(b), and 1(c). In Fig. 1, a smooth interfacial transition layer of \sim 4 monolayers was observed along the ZnSe/GaAs interface. This indicates that the transition layer is not a completely

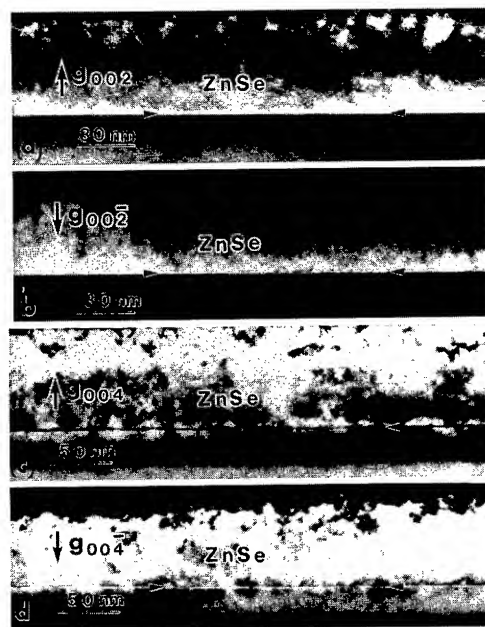


FIG. 2. TEM cross-sectional (a) 002, (b) $00\bar{2}$, (c) 004, and (d) $00\bar{4}$ dark field images from a sample with 1 min Te exposure of the GaAs surface at 400 °C prior to the growth of the ZnSe film. The Ga₂Te₃-like interfacial layer is marked by arrowheads.

intermixed layer of Ga, As, Zn, and Se atoms. No misfit dislocation or other phase with different structure was observed in TEM plan-view images, indicating that the interfacial layer is pseudomorphic to ZnSe or GaAs. In addition, the structure of the interfacial layer is coherent to the ZnSe and GaAs lattice in high resolution TEM (HRTEM) images (not shown). Thus, a zinc-blende structure with a different structure factor of the interfacial layer is assumed. In Fig. 1, the contrast of the interfacial layer remains bright under TEM dark field 002-type reflections [see Figs. 1(a) and 1(b)] and changes to dark contrast under 004-type reflections [see Fig. 1(c)]. The TEM contrast study in Fig. 1 indicates that the values of 002- and 004-type structure factors of the interfacial layers are larger and smaller than those of ZnSe film and GaAs substrate, respectively. The structure factors of 002- and 004-type reflections are given by a difference and an addition, respectively, of individual atomic scattering factors in zinc-blende structure. Thus, this observation satisfies a transition structure containing vacancies in the interfacial layers. A Ga₂Se₃-like interfacial layer has been suggested with a ZnSe film grown on Se-reacted GaAs surface.²⁷ For instance, there is no contribution to the atomic scattering at the lattice sites with Ga vacancies, therefore, a bright and dark contrast of the interfacial layer is obtained under 002- and 004-type reflections, respectively (see Fig. 1). However, we cannot determine the exact stoichiometric composition of the vacancy-contained interfacial layers.

2. Te-exposed and -reacted ZnSe/GaAs interfaces

Figure 2 shows the TEM 002- and 004-type dark field images obtained from a sample grown on Te-exposed GaAs-(2 \times 4). In Figs. 2(a) and 2(b), the contrast of the Te-exposed

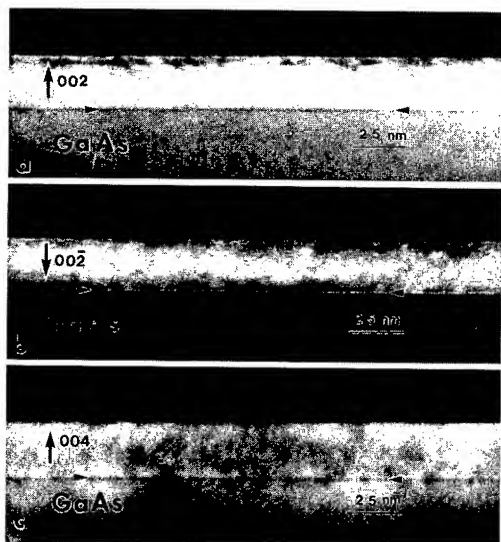


FIG. 3. TEM cross-sectional (a) 002, (b) $00\bar{2}$, and (c) 004 dark field images from a sample with 1 min Zn exposure of an As-stabilized GaAs-(2 \times 4) epilayer prior to the growth of the thin ZnSe layer. The Zn-As interfacial layer is marked by arrowheads.

interfacial layer (~ 3 –4 monolayers) remains bright under 002 and $00\bar{2}$ reflection conditions. This indicates that the interfacial layer has a value of the 002-type structure factor larger than those of GaAs and ZnSe. On the other hand, the contrast of the interfacial layer reverses sign with the change in the sign of the 004-type reflection vectors, indicating that the value of the 004 structure factor of the interfacial layer is between that of the ZnSe film and the GaAs substrate. In addition, the Te has a value of the atomic scattering factor almost as twice as large as those of the Ga, As, Zn, or Se atoms.²⁸ Thus, an interfacial layer containing Te atoms and Ga vacancies satisfies the observation in Fig. 2 of the TEM dark field images. In this case, the difference between individual atomic scattering factor is increased due to either a high atomic scattering factor of Te atoms or no contribution in atomic scattering at the lattice sites with Ga vacancies. This gives rise to an increase on the 002 structure factor. Consequently, a bright contrast of the interfacial layer is observed under 002-type reflections as shown in Figs. 2(a) and 2(b). On the other hand, the value of the 004 structure factor is increased by Te atoms and decreased by Ga vacancies, so that the value of the 004 structure factor of the interfacial layer is between that of GaAs and ZnSe as shown in Figs. 2(c) and 2(d). Therefore, an interfacial layer of ~ 3 –4 monolayers containing Ga₂Te₃-like chemical composition is suggested from Fig. 2.

3. Interfacial layers in ZnSe/GaAs with Zn treatment on an As-stabilized GaAs-(2 \times 4) surface

Figure 3 shows TEM (a) 002, (b) $00\bar{2}$, and (c) 004 dark field images for a sample with 1 min Zn treatment of an As-stabilized GaAs-(2 \times 4). Under the 002-type reflections, a sharp interfacial transition layer of ~ 2 monolayers was observed along the ZnSe/GaAs interface. In Figs. 3(a) and 3(b),



FIG. 4. TEM cross-sectional (002)-type dark field images of ZnSe/GaAs with films grown on Zn-exposed GaAs-c(4 \times 4) (a) without, (b) with 10 min at 400 °C, and (c) with 30 min at 440 °C of thermal annealing.

the contrast of the interfacial layer reverses sign with the change in the sign of the 002-type reflection vectors, indicating that the value of the 002 structure factor of the interfacial layer is between that of the ZnSe film and the GaAs substrate. In addition, under 004-type reflections of TEM dark field images, this interfacial layer remains in dark contrast indicating a smaller 004 structure factor of the interfacial layer than that of either ZnSe film or GaAs substrate [see Fig. 3(c)]. Therefore, the interfacial layer observed in Fig. 3 has a 002 structure factor between that of ZnSe and GaAs and a smaller 004 structure factor than that of either ZnSe film or GaAs substrate. In this fashion, by comparing the atomic scattering factors among Ga, As, Zn, and Se, an interfacial layer containing a majority of Zn-As interfacial layer of ~ 2 monolayers is suggested by the TEM observation. In contrast to the samples grown on Se-exposed GaAs-(4 \times 6) with 4 monolayers of Ga₂Se₃-like interfacial layers, there is no significant atomic interdiffusion in samples grown on Zn-exposed GaAs-(2 \times 4). From a previous study of Zn preadsorption on GaAs-(2 \times 4) at 290 °C,²⁹ however, the Zn coverage on the order of 0.1 monolayer rather than a full Zn termination is suggested by x-ray photoemission spectroscopy.

4. Interfacial layers in ZnSe/GaAs with Zn treatment on the As-rich GaAs-c(4 \times 4) with and without thermal annealing

Figure 4(a) shows a TEM 002 dark field image of ZnSe/GaAs with 1 min Zn exposure on an As-rich GaAs-c(4 \times 4) prior to the growth of the ZnSe film. The same bright and dark diffraction contrasts of the interfacial layer were obtained under 002- and 004-type reflections for samples grown on Zn-exposed GaAs-(2 \times 4) and -c(4 \times 4). However, a thicker Zn-As interfacial layer of ~ 3 –4 monolayers was obtained in samples grown on Zn-exposed GaAs-c(4 \times 4) as

shown in Fig. 4(a). The surface of the As-stabilized GaAs-(2×4) and -c(4×4) epilayers are known to exist as single and double As-terminated layers, respectively.²⁵ In addition, the c(4×4) surface reconstruction became (2×2)-like after 1 min of Zn exposure at 260 °C by RHEED observation. This indicates that Zn indeed reacted with the As-rich GaAs-c(4×4) and the ordered arrangement of the c(4×4) reconstruction had changed due to Zn exposure. This is consistent with the observation in Fig. 4(a) that Zn may interact with both top and bottom As layers on the surface of Zn-exposed GaAs-c(4×4) to form ~3–4 monolayers of Zn–As interfacial layer. Figures 4(b) and 4(c) show TEM 002 and 002̄ dark field images obtained from ZnSe/GaAs samples with 10 min annealing at 400 °C and 30 min annealing at 440 °C on the Zn-exposed GaAs-c(4×4), respectively. In Fig. 4(b), a rough and diffusive contrast of the ZnSe/GaAs interfacial layer is obtained from a sample with 10 min of thermal annealing at 400 °C on the Zn-exposed GaAs-c(4×4). Similar rough contrast of the interfacial layer was obtained under 002̄- and 004-type reflection conditions. The thickness of the Zn–As interfacial layer shown in Fig. 4(a) was reduced and became rough after 10 min thermal annealing at 400 °C. This suggests that the interaction between Zn and As atoms is reduced; however, local clustering of the Zn–As interactive layers was taking place upon thermal annealing.

On the other hand, very smooth and no clear contrast of the interfacial layer was obtained in a sample grown on a Zn-exposed GaAs-c(4×4) with 30 min of thermal annealing at 440 °C as shown in Fig. 4(c). The surface reconstruction of the GaAs surface after thermal annealing was (2×4) prior to the growth of the thin ZnSe film. In this case, a weaker contrast and much thinner thickness (~0–2 monolayers) of the interfacial layer is obtained than that of the interfacial layer obtained in samples grown on Zn-exposed GaAs-(2×4) and -c(4×4) without thermal annealing. This indicates that the chemical composition of the interfacial layer in Fig. 4(c) is a relatively intermixed layer containing Ga, As, Zn, and Se atoms. Similar diffraction contrasts were obtained in samples with different annealing times of ~10–70 min at 440 °C.

As shown in Fig. 4, the thicknesses of the Zn–As interactive interfacial layers decreased due to thermal annealing on the Zn-exposed GaAs-c(4×4). The c(4×4) surface reconstruction of GaAs epilayer became (2×2)-like with 1 min of Zn exposure at 260 °C by RHEED observation. After thermal annealing, the surface reconstructions of GaAs-c(4×4) epilayers were changed to (2×3)- and then (2×4)-like under Zn exposure with annealing temperature at 400 and 440 °C, respectively. In this fashion, the (2×3)-like surface reconstruction is likely to be a transitional surface structure from the Zn-exposed As-rich GaAs-c(4×4) to GaAs-(2×4) during the desorption of surface excess As atoms by thermal annealing. Therefore, the results in Figs. 4(a)–4(c) suggest that the interaction between Zn and the surface of GaAs-c(4×4) is decreased by thermal annealing, probably by decreasing the excess As coverage on the GaAs-c(4×4) surface.

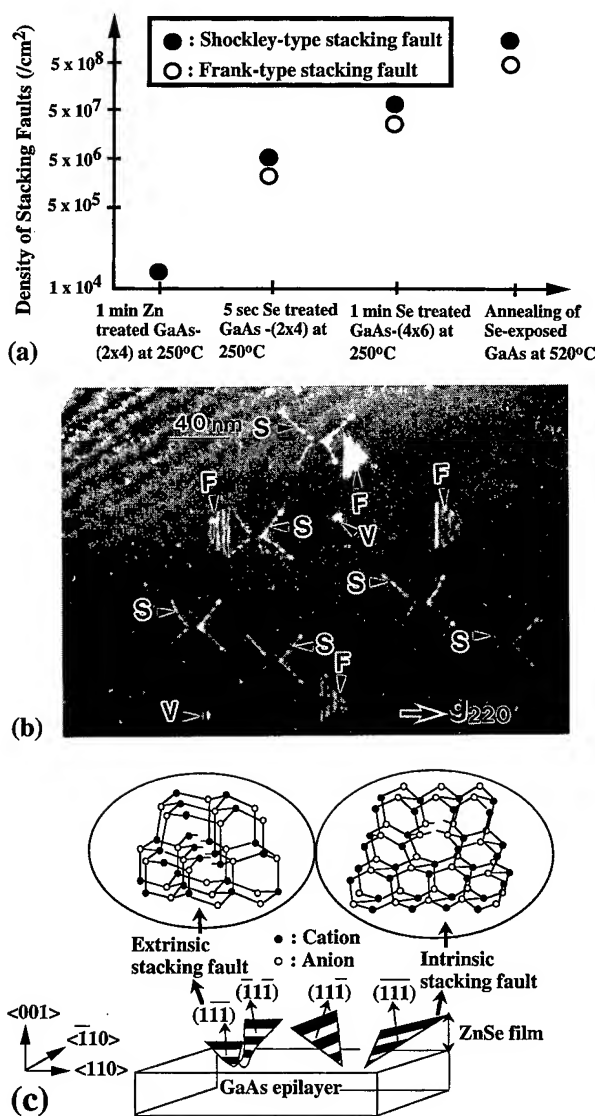


FIG. 5. (a) Dependence of density of stacking faults on Se contamination or interaction with the GaAs surface. (b) TEM plan view 220 weak beam image of a sample with thermal annealing at 520 °C on a Se-exposed GaAs-(4×6). (c) A schematic diagram of the intrinsic- and extrinsic stacking faults. In (b) the vacancy loops are labeled V and the Shockley- and Frank-type stacking faults are labeled S and F, respectively.

B. Distribution and structure of stacking faults in ZnSe grown on Se-exposed and -reacted GaAs

1. Distribution of stacking faults

Figure 5(a) shows a distribution of stacking faults in ZnSe films grown on Se-exposed GaAs surfaces with and without thermal annealing as indicated in the figure. In this figure, the horizontal axis represents different degrees of interaction between Se and GaAs surfaces. In Fig. 5(a), the densities of Shockley- and Frank-type stacking faults are of ~5×10⁷/cm² in samples with 1 min Se exposure of a Ga-rich GaAs-(4×6) surface. In addition, very high densities of Frank-type vacancy loops, Frank- and Shockley-type stacking faults of ~5×10⁸–1×10⁹/cm² were generated as shown in Fig. 5(b) of a TEM plan-view image including the ZnSe/GaAs inter-

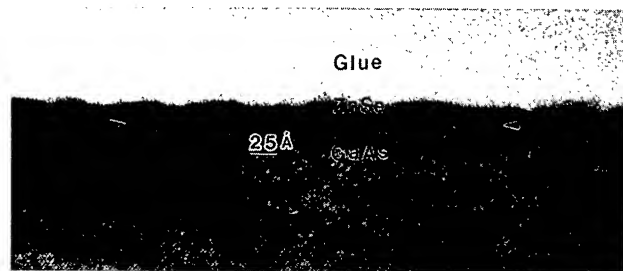


FIG. 6. HR TEM image of a very thin ZnSe film (~ 5 –9 monolayers) grown on Se-exposed GaAs-(4×6). In this figure, the island growth mode of the ZnSe film is clearly recognized.

face. In this case, thermal annealing at 520 °C was applied to the Se-exposed Ga-rich GaAs-(4×6) surface to desorb As atoms near the GaAs surface and increase the degree of interaction between Se and GaAs surface. TEM 002- and 004-type reflections were used to study the chemical composition of the ZnSe/GaAs interfaces. No difference in contrast and thickness of the interfacial layer of ~ 2 monolayers from TEM observation was recognized in samples with 1 min Zn and 5 s Se exposure on the As-stabilized GaAs-(2×4) surfaces (see Fig. 3). However, a Ga_2Se_3 -like interfacial layer of ~ 4 monolayers was identified in samples grown on Se-exposed GaAs-(4×6) surfaces with and without thermal annealing (see Fig. 1). In contrast to the low defect density in samples grown on Zn-exposed As-stabilized GaAs-(2×4) surfaces, the densities of Shockley- and Frank-type stacking faults increase as the degree of interaction between Se and GaAs surface increases as shown in Fig. 5(a). In samples grown on the Se-exposed or reacted- GaAs-(2×1), spotty RHEED reconstructions were observed at the initial stages of ZnSe growth. This indicates that island growth mode of ZnSe films was taking place on the (2×1) Se-terminated GaAs surfaces. The island growth mode of ZnSe films is clearly observed (as shown in Fig. 6) as a cross-sectional HRTEM image. In Fig. 6, ~ 5 –9 monolayers of ZnSe film is grown on the Se-exposed GaAs-(4×6) at a growth stage upon island coalescence. However, no stacking faults or other defects is observed at the boundaries among the ZnSe growth islands.

2. Structure of Frank- and Shockley-type stacking faults from ZnSe/GaAs interface

The glissile Shockley partial dislocations have Burgers vector on the $\{111\}$ -type slip planes of zinc-blende lattice. Thus, the Shockley partial dislocations can glide easily and become curved [see Fig. 5(b)] on the faulted plane of the stacking faults probably due to the lattice strain in ZnSe/GaAs. On the other hand, the sessile Frank partial has a Burgers vector perpendicular to the $\{111\}$ -type slip planes, thus, the Frank partials cannot glide on the $\{111\}$ -type faulted planes. Therefore, the Frank-type stacking faults are bound by a straight dislocation line as shown in Fig. 5(b). In Fig. 5(b), the stacking faults bound by Shockley partial dislocations are labeled S. The fault planes of the Shockley-type stacking faults extend to the film's surface intersecting it by

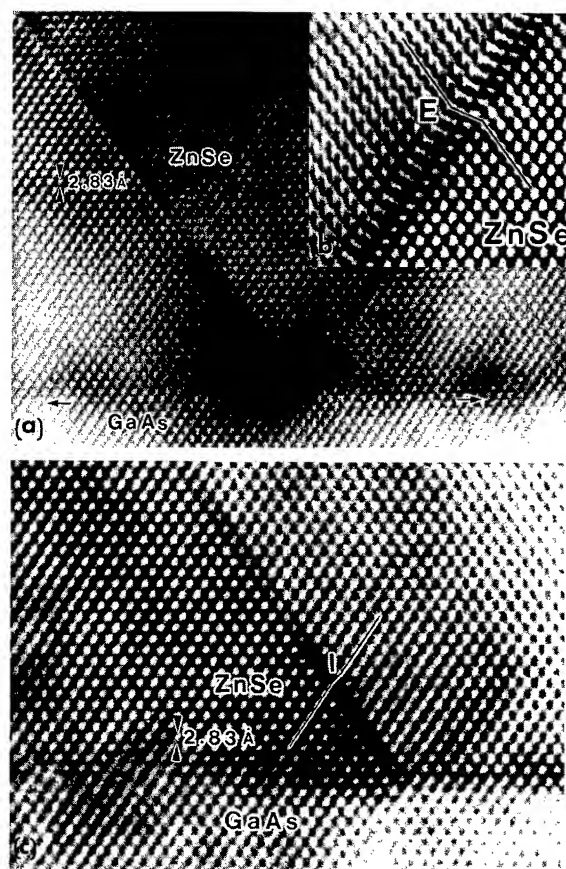


FIG. 7. (a), (b) HRTEM images of an extrinsic- (labeled E) and (c) an intrinsic-type (labeled I) stacking faults obtained from a sample with a film grown on a Se-reacted GaAs surface. (a) and (b) were observed with electron beam along $\langle 110 \rangle$ and (c) was obtained along $\langle \bar{1}10 \rangle$.

lines parallel to $\langle 110 \rangle$ direction as recognized by the elongated direction of etch pits from the back side of the TEM specimen. The nature and structure of the Shockley-type stacking fault is studied in the lattice resolution using HRTEM observed along the $\langle 110 \rangle$ irradiated direction of electron beam as shown in Figs. 7(a) and 7(b). In Fig. 7(a), the Shockley-type stacking fault originates from the ZnSe/GaAs interface and extends throughout the ZnSe film. In Fig. 7(b) of an enlarged HRTEM image of the stacking fault E, the extrinsic nature of the fault is clearly recognized with an additional $(\bar{1}\bar{1}\bar{1})$ or (111) lattice plane inserted between other close-packed $(11\bar{1})$ or $(\bar{1}11)$ lattice planes in the ZnSe film. Formation of an extrinsic-type stacking fault in the zinc-blende structure involves the creation of a pair of $1/2\{111\}$ lattice planes (cation and anion lattice planes). This is recognized as shown in Fig. 7(b) of the extrinsic-type defects with a pair of additional $1/2(\bar{1}\bar{1}\bar{1})$ interference lattice fringe at the stacking faults E.

Other kinds of stacking faults [labeled F in Fig. 5(b)] obtained in samples grown on Se-exposed or -reacted surfaces of GaAs epilayers, were recognized as intrinsic Frank-type stacking faults lying on $(11\bar{1})$ or $(\bar{1}11)$ planes by TEM bright field imaging of plan-view specimens and HRTEM observation. The fault planes of the Frank-type stacking

faults extend to the film's surface intersecting it by lines parallel to the $\langle 110 \rangle$ direction [see Fig. 5(b)]. Figure 7(c) is a $\langle 110 \rangle$ HRTEM image of the Frank-type stacking faults [labeled F in Fig. 5(b)] observed along the $\langle 110 \rangle$ irradiated direction of the electron beam in TEM. In this figure, the stacking fault originates from the ZnSe/GaAs interface and the intrinsic nature of the Frank-type stacking fault is recognized with a slight decrease in spacing of the lattice fringe at the stacking fault I in Fig. 7(c). The density of Frank-type stacking faults is of $\sim 5 \times 10^7/\text{cm}^2$ and $\geq 5 \times 10^8/\text{cm}^2$ in samples grown on Se-exposed Ga-rich GaAs-(4 \times 6) surfaces with and without thermal annealing, respectively. Thus, the intrinsic-type stacking fault originates from the ZnSe/GaAs interface by removing a $\langle 111 \rangle$ or $\langle \bar{1}\bar{1}\bar{1} \rangle$ plane containing Zn and Se atoms.

Therefore, the Shockley-type stacking faults are identified as extrinsic-type defects lying on $\langle 111 \rangle$ and $\langle \bar{1}\bar{1}\bar{1} \rangle$ planes. Since the zinc-blende structure is noncentrosymmetric, the extrinsic-type stacking faults lying on $\langle 111 \rangle$ and $\langle \bar{1}\bar{1}\bar{1} \rangle$ planes are bound by partial edge dislocations with a core structure terminated at additional cations. On the other hand, the intrinsic Frank-type stacking faults lying on $\langle 111 \rangle$ or $\langle \bar{1}\bar{1}\bar{1} \rangle$ are bound by partial edge dislocations with a core structure terminated on anions. A schematic diagram of the structures of the extrinsic- [labeled S in Fig. 5(b)] and intrinsic-type [labeled F in Fig. 5(b)] stacking faults bound by partial edge dislocations is proposed in Fig. 5(c).²² In this figure, the extrinsic-type Shockley partials have a core structure with additional cations (such as cation interstitials), and the intrinsic Frank partials have a core structure with a lack of cations (such as cation vacancies).

3. Vacancy-loops generated due to thermal annealing

Furthermore, a high density of tiny defects of $\geq 10^9/\text{cm}^2$ [labeled V in Fig. 5(b)] was generated in samples grown on Se-exposed Ga-rich GaAs-(4 \times 6) surfaces with thermal annealing at 520 °C. However, no such kind of tiny defect was found in samples grown on Se-exposed GaAs-(4 \times 6) without thermal annealing. In TEM dark field images (not shown), the diffraction vector points toward a bright lobe of these tiny defects, indicating the intrinsic characteristic of these tiny Frank-type loops.³⁰ Figures 8(a) and 8(b) show a HRTEM image and a schematic diagram of a vacancy loop obtained in a sample grown on a Se-reacted Ga-rich GaAs-(4 \times 6) surface with thermal annealing at 520 °C. In Fig. 8(a), the tiny defect originates at an area near the ZnSe/GaAs interface and extends along $\{111\}$ -type planes with dimension of ~ 10 – 12 nm within the ZnSe film. The intrinsic nature of the vacancy loops is recognized by the positions of horizontal $\langle 002 \rangle$ atomic lattice fringes in Fig. 8(a). In this figure, the $\langle 002 \rangle$ atomic lattice fringes between the fault planes of the vacancy loop have positions slightly below the positions of the $\langle 002 \rangle$ lattice fringes near the vacancy loop as indicated by the double-head arrows. A schematic diagram of the vacancy loop observed in Fig. 8(a) is shown in Fig. 8(b) with missing atoms on $\{111\}$ planes.

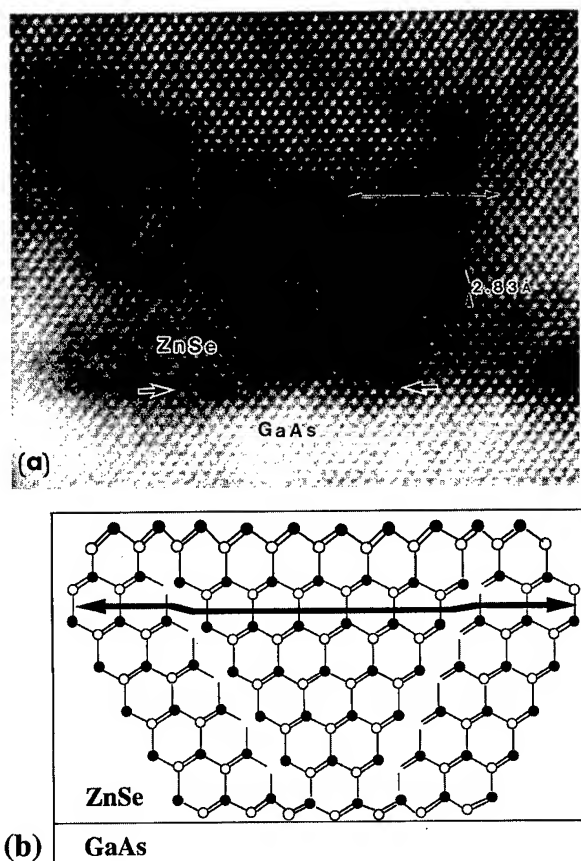


FIG. 8. (a) HRTEM image of a vacancy loop originating from the ZnSe/GaAs interface. (b) A schematic diagram of the vacancy loop shown in (a). The intrinsic nature of the vacancy loop is shown by the double-head arrows in (a) and (b).

C. Distribution and structure of stacking faults in ZnSe grown on Zn-exposed As-terminated GaAs

1. Distribution of stacking faults and precipitates from ZnSe/GaAs interface

Figure 9(a) shows the distribution of stacking faults in ZnSe films as a function of surface stoichiometry of As-terminated GaAs surface. In this figure, a very low density of stacking faults in the range of $\sim 10^4/\text{cm}^2$ was obtained in samples with 1 min Zn treatment on an As-stabilized GaAs-(2 \times 4) surface. However, the densities of stacking faults and precipitates increase as the surface coverage of c(4 \times 4) reconstruction increased on the As-stabilized GaAs-(2 \times 4) surface. Figure 9(b) show a TEM plan-view image obtained from a sample with a ZnSe film grown on an As-rich GaAs-c(4 \times 4) surface. In contrast to the low defect densities in samples with ZnSe films grown on As-stabilized GaAs-(2 \times 4) epilayers, a very high density of tiny defects (12–15 nm) of $\geq 5 \times 10^9/\text{cm}^2$ [labeled P in Fig. 9(b)] was obtained in samples grown on an As-rich GaAs-c(4 \times 4) surface [see Fig. 9(a)]. In addition, a high density of stacking faults bound by Shockley partials of $\sim 1 \times 10^9/\text{cm}^2$ [labeled S in Fig. 9(b)] was obtained in the samples grown on the As-rich GaAs-c(4 \times 4) surfaces.

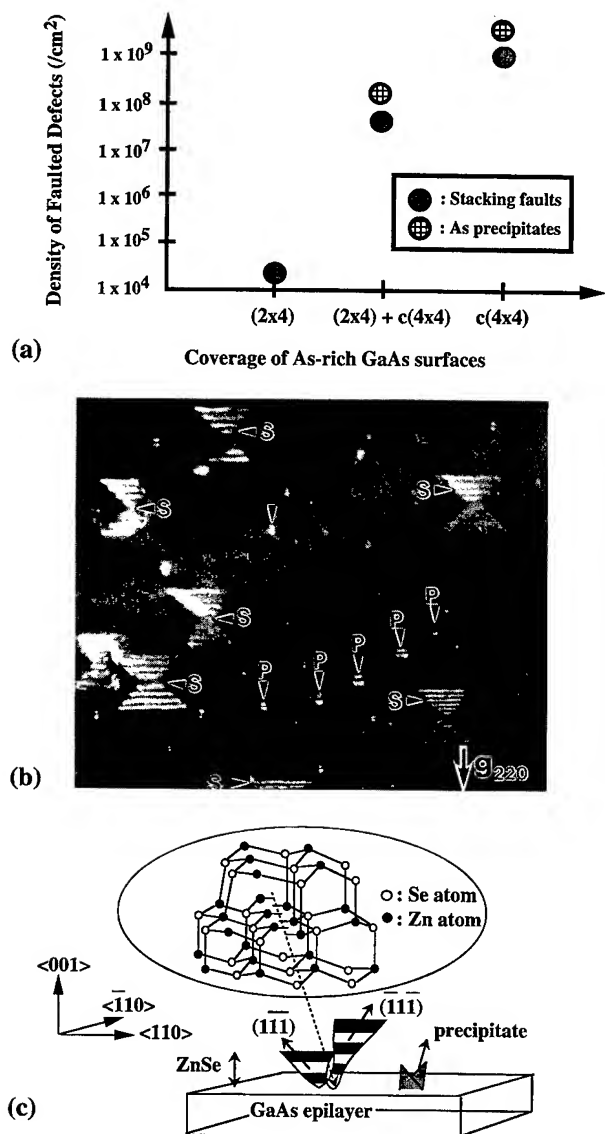


FIG. 9. (a) Dependence of the densities of As precipitates and Shockley-type stacking faults on the coverage of As-terminated GaAs surface. Zn passivation was applied to the As-terminated GaAs surfaces prior to the growth of the films. (b) TEM 220 weak beam image and (c) a schematic diagram of the defects obtained from a sample with Zn exposure on a As-rich GaAs-c(4×4) surface.

2. Structure of stacking faults and precipitates

As the stacking faults are labeled S in Fig. 5(b) and E in Fig. 7(b), the Shockley-type stacking faults S in Fig. 9(b) are intrinsic defects and originate from the ZnSe/GaAs interface. Thus, these stacking faults are identified as extrinsic-type defects lying on (111) and $(\bar{1}\bar{1}\bar{1})$ planes. On the other hand, a high density of tiny defects [labeled P in Fig. 9(b)] was generated in samples grown on Zn-exposed GaAs-c(4×4). In TEM 220-type plan-view images [see Fig. 9(b)], there is a line of zero contrast crossing these tiny defect with a direction normal to the operated $\langle 220 \rangle$ or $\langle \bar{2}\bar{2}0 \rangle$ two-beam. Thus, these tiny defects have a symmetrical strain field (or structure) with respect to the $\langle 110 \rangle$ and $\langle \bar{1}\bar{1}0 \rangle$ directions. A sche-

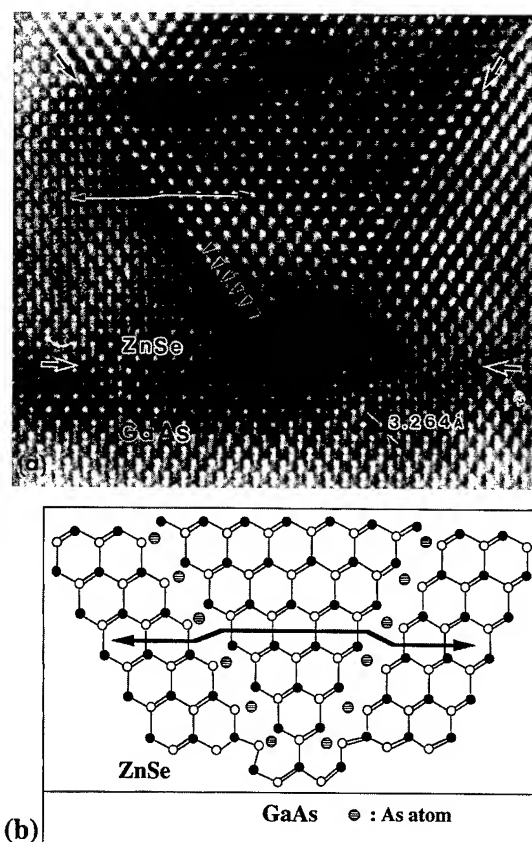


FIG. 10. (a) HRTEM image of a precipitate originating from the ZnSe/GaAs interface. The precipitation is marked by arrowheads with an additional row of lattice fringes inserted between the $1/2\{111\}$ interference fringes. (b) A schematic diagram of the precipitate shown in (a). The extrinsic nature of the precipitate is shown by the double-head arrows in (a) and (b).

matic diagram of the crystalline defects generated in samples with 1 min Zn exposure of an As-rich GaAs-(4×4) surface is proposed as shown in Fig. 9(c).²²

3. Atomic lattice resolution of the As precipitates

Figures 10(a) and 10(b) show a HRTEM image and a schematic diagram of a tiny defect labeled P in Fig. 9(b). In Fig. 10(a), the atomic lattice fringes with half the period of the $\{111\}$ planes are clearly resolved at and near the tiny defect. This additional intensity variation in the lattice image is induced by the interference between the $\langle 1\bar{1}\bar{1} \rangle$ and $\langle \bar{1}\bar{1}1 \rangle$ diffracted beams.³¹ The interference $1/2\{111\}$ lattice fringes may form due to a difference of extinction distances or strain field at areas around the tiny defect as shown in Fig. 10(a). These $1/2\{111\}$ lattice fringes allow for the possibility of resolving the defect structure with a space resolution corresponding to individual $1/2\{111\}$ cation and anion lattice planes. In Fig. 10(a) the tiny defect also originates from the ZnSe/GaAs interface with an additional row of $1/2\{111\}$ atomic lattice fringe inserted between different $1/2\{111\}$ planes, as indicated by the arrowheads shown in Fig. 10(a). This indicates the extrinsic nature of the tiny defect which is formed by inserting only an extra row of anions or cations, but not in both. A schematic diagram of the tiny defect is

shown in Fig. 10(b). Due to the additional atoms inserted along the $\{111\}$ -type planes (precipitation), the horizontal (002) atomic lattice fringes between the fault planes of the tiny defect (precipitate) have positions slightly above the positions of the (002) lattice fringes near the precipitate, as indicated by the double-head arrows shown in Fig. 10.

The extrinsic tiny defects shown in Figs. 9(b) and 10 are likely to be As precipitates due to the local nonstoichiometric arrangement near the ZnSe/GaAs interface:

(1) The precipitates are extrinsic Frank-type dislocation loops (see Fig. 10) with an additional $1/2\{111\}$ plane inserted in the film and located near the ZnSe/GaAs interface. Thus, they are formed due to clustering and interdiffusion of excess atoms from ZnSe/GaAs interface after the growth of the ZnSe films.

(ii) These tiny defects were observed only in samples with 1 min Zn exposure of a GaAs- $c(4\times4)$, but not a GaAs- (2×4) , As-rich GaAs surface [see Fig. 9(a)]. In addition, the density of precipitates increases as the surface As coverage of $c(4\times4)$ reconstruction increases. The surfaces of the As-rich GaAs- (2×4) and $-c(4\times4)$ epilayers are known to exist as single and double As-stabilized layers, respectively.²⁵

(iii) The density of the precipitates decreased due to thermal annealing at 440 °C (described in Sec. III D). In this case, excess As atoms on GaAs- $c(4\times4)$ surface may be desorbed during thermal annealing (see Fig. 4).

(iv) Furthermore, only extrinsic Shockley-type stacking faults, but not the precipitates, are observed in the films grown on Zn-exposed As-stabilized GaAs- (2×4) under Zn- and Se-rich growth conditions.¹⁷ This indicates that excess Zn and Se on the GaAs surface under Zn- and Se-rich growth conditions, respectively, may contribute to the formation of stacking faults extending throughout the films, but not the Frank-type dislocation loops (precipitates) buried near the ZnSe/GaAs interface.

Therefore, it is likely that the tiny defects (Frank-type dislocation loops) near the ZnSe/GaAs interface, as shown in Figs. 9(b) and 10, are mainly As precipitates formed due to clustering and diffusion of excess As atoms from ZnSe/GaAs interface after the growth of the films on the Zn-exposed GaAs- $c(4\times4)$. The size of the As precipitates is of ~ 15 nm lying on $\{111\}$ -type planes. If we assume that these As precipitates with a zinc-blende stacking (see Fig. 10) have a lattice constant close to that of ZnSe, there are $\sim 5 \times 10^3$ As atoms per precipitate. The density of the As precipitates is in the range of $\geq 10^{10}/\text{cm}^2$ in samples grown on Zn-exposed GaAs- $c(4\times4)$. Thus, there are $\geq 5 \times 10^{13}$ As atoms/ cm^2 diffusing from the interface after the growth of the ZnSe film on the Zn-exposed GaAs- $c(4\times4)$ epilayer. The atomic density on (001) ZnSe or GaAs surface is of $\sim 5 \times 10^{14}$ atoms/ cm^2 . There are approximately ≥ 0.1 monolayer of As atoms which contribute to the formation of As precipitates.

D. Dependence of defect density in ZnSe/GaAs on thermal annealing

1. Effect of thermal annealing on defect generation in ZnSe grown on Te-exposed GaAs- (2×4)

Figure 11 shows a distribution of defects obtained in

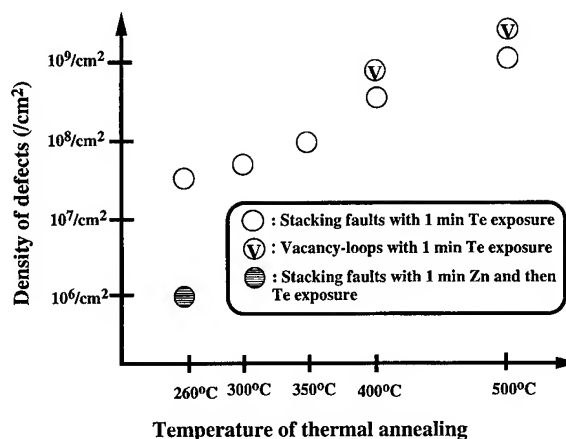


FIG. 11. Distribution (or densities) of stacking faults in samples grown on Te-exposed GaAs with thermal annealing at different temperature (260–500 °C). A low density of fault defects obtained in a sample grown on a 1 min Zn- and then 1 min Te-exposed GaAs- (2×4) at 260 °C is added in this figure for comparison.

samples grown on Te-exposed GaAs- (2×4) with and without thermal annealing. Very streaky RHEED intensity oscillations were observed right after the growth of the ZnSe films on the Te-exposed GaAs.²⁶ This indicates that relatively layer-by-layer growth of ZnSe film was taking place on the Te-exposed GaAs. However, both Frank- and Shockley-type stacking faults with defect densities of $\sim 5 \times 10^7$ – $1 \times 10^9/\text{cm}^2$ were observed in samples grown on Te-exposed GaAs with and without thermal annealing (see Fig. 11). In Fig. 11, the density of the stacking faults in ZnSe/GaAs increases as the temperature of thermal annealing on the Te-exposed GaAs increased. In addition, a high density of vacancy loops was generated in samples grown on 1–2 min thermal annealing at 400 and 500 °C of the Te-exposed GaAs. In Fig. 11, the density of the defects increases as the interaction between Te atoms and GaAs surface increases by increasing the annealing temperature. However, the interaction between Te atoms and the GaAs surface was prevented by Zn treatment on the GaAs- (2×4) at 260 °C prior to the Te exposure as shown in Fig. 11. In this case, the density of stacking faults is in the range of $\leq 10^6/\text{cm}^2$ with 1 min of Zn and then Te exposed on the GaAs- (2×4) prior to the growth of the ZnSe films. Therefore, the generation of the fault defect is greatly dependent on the interaction between group II (Se and Te) atoms and the GaAs surface as shown in Figs. 5(a) and 11. This interaction can be greatly prevented by Zn treatment on the GaAs- (2×4) prior to the growth of the films with a very low density of the fault defects [see Figs. 9(a) and 11].

2. Effect of surface thermal annealing on defect generation in ZnSe grown on Zn-exposed GaAs- $c(4\times4)$

Figure 12 shows a distribution of stacking faults in samples grown on Zn-exposed GaAs- $c(4\times4)$ without and with thermal annealing at 400 and 440 °C. The Zn shutter was opened during the thermal annealing to prevent possible

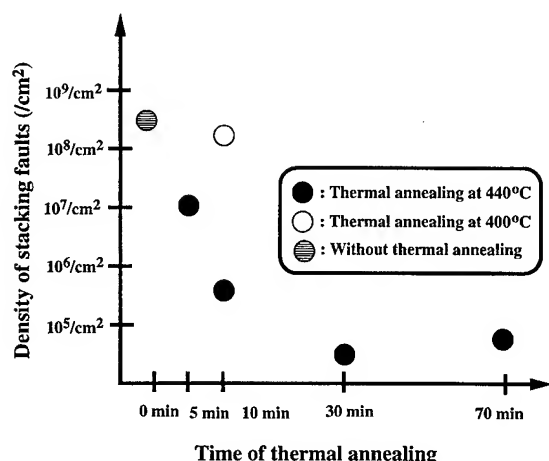


Fig. 12. Distribution (or densities) of stacking faults in samples grown on Zn-exposed GaAs-c(4×4) with 10 min of thermal annealing at 400 °C and 5–70 min of thermal annealing at 440 °C. Also, the defect density in a sample grown on Zn-exposed GaAs-c(4×4) at 260 °C without any thermal annealing is added in the figure for comparison.

contamination by Se atoms. In this figure, high densities of stacking faults ($\geq 5 \times 10^8/\text{cm}^2$) and As precipitates ($\geq 5 \times 10^9/\text{cm}^2$) were generated in samples grown on Zn-exposed GaAs-c(4×4) without thermal annealing. The densities of stacking faults and precipitates decreased in samples grown on Zn-exposed GaAs-c(4×4) with 10 min thermal annealing at 400 °C (see Fig. 12). Furthermore, the density of the stacking faults is greatly reduced in samples with 5–70 min of thermal annealing under Zn exposure on the Zn-exposed GaAs-c(4×4) prior to the growth of the ZnSe films. In these samples, no precipitates or vacancies were found by TEM and the density of the As precipitates is very low. It is known that the desorption temperature of excess As on GaAs-c(4×4) surface is of ≥ 400 °C.¹⁸ In our case, the densities of the stacking faults and precipitates were greatly reduced due to thermal annealing at 440 °C on the Zn-exposed As-rich GaAs-c(4×4). In addition, the results in Figs. 4(a)–4(c) suggest that the interaction between Zn and the surface of GaAs-c(4×4) and the thickness of the Zn–As interactive layer is decreased by thermal annealing at 440 and 440 °C probably by decreasing the As coverage on the GaAs surface [see Figs. 4(a)–4(c)]. Therefore, these results suggest that the density of the crystalline defects is greatly dependent on the interaction between Zn and excess As atoms on the GaAs surface.

E. Interface chemistry vs defect generation

1. Defect generation in ZnSe/GaAs due to Se or Te interaction with GaAs surface

As shown in Figs. 5(a), both intrinsic- and extrinsic-type fault defects with densities of $\sim 5 \times 10^7/\text{cm}^2$ were generated from the ZnSe/GaAs interfaces in samples grown on Se-exposed Ga-rich GaAs-(4×6) surfaces. In addition, high densities of vacancy loops ($\geq 1 \times 10^9/\text{cm}^2$), intrinsic Frank-type ($\sim 5 \times 10^8/\text{cm}^2$), and extrinsic Shockley-type stacking faults ($\sim 1 \times 10^9/\text{cm}^2$) were generated in the samples with

1–2 min of annealing at 520 °C of a Se-exposed Ga-rich GaAs-(4×6) [see Fig. 5(a)]. Therefore, thermal annealing at 520 °C induced additional vacancies (vacancy loops and intrinsic-type stacking faults) and extrinsic-type stacking faults with cation-terminated dislocation cores (see Fig. 5). In these samples, an ~ 4 monolayers of Ga₂Se₃-like interfacial layer was obtained as shown in Fig. 1. At the initial stages of growth, an island growth mode of the ZnSe film took place on the Se-exposed or -reacted GaAs surface suggested by RHEED observation and HRTEM imaging of the ZnSe/GaAs as shown in Fig. 6. On the other hand, similar distribution and the same kind of fault defects were also obtained in samples grown on Te-exposed GaAs-(2×4) with and without thermal annealing (see Fig. 11). The densities of Frank- and Shockley-type stacking faults are of $\sim 5 \times 10^7/\text{cm}^2$ in both samples grown on Se- and Te-exposed GaAs surfaces at 260 °C and the densities of intrinsic- and extrinsic-type stacking faults increase as a function of annealing temperature. In both cases, 3–4 monolayers of Ga₂Se₃- and Ga₂Te₃-like interfacial layers are found in samples grown on Se- and Te-exposed GaAs surfaces, respectively (see Figs. 1 and 2). However, in contrast to the island growth mode of ZnSe grown on Se-exposed GaAs-(4×6) (see Fig. 6), a layer-by-layer growth mode of the ZnSe films took place with films grown on Te-exposed GaAs surfaces. Therefore, these results suggest that the generation of the fault defect in ZnSe/GaAs is greatly dependent on the interaction between Se or Te atoms with the GaAs surface. In addition, the growth mode of ZnSe film may not have a direct correlation with the generation of fault defects in ZnSe/GaAs.

In samples without thermal annealing, formation of Ga₂Se₃- or Ga₂Te₃-like interfacial layers may be associated with the replacement of As by Se or Te interdiffusion near GaAs surface to form 3–4 monolayers of Ga–Se or Ga–Te interfacial layers with Ga vacancies in the zinc-blende structure (see Figs. 1 and 2). In both cases, the densities of intrinsic- and extrinsic-type defects increased due to thermal annealing (see Figs. 5 and 11). In annealed samples, the generation of additional Ga vacancies may be associated with the formation of additional Ga interstitials on the GaAs surface. Unlike the As atoms on the GaAs surface, the surface Ga interstitials may not desorb easily from the GaAs surface even at the annealing temperature of ~ 520 °C. Stacking errors due to excess Ga interstitials on the GaAs surface may have provided the nucleation sites of extrinsic-type stacking faults after the growth of ZnSe film. Extrinsic and intrinsic stacking faults can be generated from interstitials and vacancies on GaAs surface, respectively, by repeating the same stacking errors on these point defects on different {111} planes during the growth of ZnSe films. Therefore, the generation of extrinsic cation-terminated Shockley and intrinsic anion-terminated Frank partials may be related to the formation of Ga interstitials and vacancies, respectively, due to the interaction between Se or Te atoms and the GaAs surface [see Figs. 5(a) and 11]. This speculation is supported by thermal annealing of the Se- and Te-exposed GaAs. The interaction between Ga and Se or Te (or replacement of As by

Se or Te atoms) and desorption of surface As atoms is enhanced by thermal annealing. This induces an increase on the nucleation sites (cation vacancies and interstitials) for the generation of high densities of vacancy loops, intrinsic Frank-, and extrinsic Shockley-type stacking faults [see Fig. 5(a) and 11].

2. Defect generation in ZnSe/GaAs due to Zn exposure on the GaAs-c(4×4) surface

A majority of a Zn-As interfacial layer was obtained in samples with 1 min Zn exposure on both (2×4) and c(4×4) As-stabilized GaAs surfaces [see Figs. 3 and 4(a)]. However, ~ 2 and 3–4 monolayers of Zn-As interfacial layers were obtained in samples grown on Zn-exposed GaAs-(2×4) and -c(4×4), respectively. In Fig. 9(a), the density of precipitates increases as the increase on surface As coverage of c(4×4) reconstruction on the As-stabilized GaAs-(2×4) surface and associates with an increase in the density of extrinsic-type stacking faults bound by partial edge dislocations with a core structure terminated on additional cations (such as Zn interstitials). In addition, the thickness of the Zn-As interfacial layers is decreased due to thermal annealing at 400 and 440 °C on the Zn-exposed GaAs-c(4×4) (see Fig. 4). The densities of stacking faults and precipitates decreased as the thickness of this interactive layer decreased by thermal annealing [see Figs. 4 and 9(a)]. Therefore, it is likely that the density of stacking faults and As precipitates increased as the interaction between Zn and excess As atoms increased.

Due to the different nature in structure between the Frank-type tiny precipitates and Shockley-type stacking faults, we believed that the origins of nucleating these defects are different. The As precipitates were inclined to GaAs and buried in ZnSe as shown in Figs. 9(b) and 10, thus, we believed that they were introduced by interdiffusion of excess As along {111} planes from the ZnSe/GaAs interface during the growth of the ZnSe film. Therefore, it is likely that local clustering of As atoms and formation of Zn interstitials took place on a Zn-exposed As-excess GaAs-c(4×4) surface. These local nonstoichiometric arrangements on the GaAs surface may provide the nucleation sites of precipitates and extrinsic-type stacking faults. The interaction between Zn and excess As (clustering of As atoms and formation of Zn interstitials) is reduced by thermal annealing with a decrease in the amount of excess As on the GaAs-c(4×4) surface (see Figs. 4 and 12). In contrast to the high defect density in samples grown on annealed Se-exposed GaAs-(4×6), the density of stacking faults and precipitates is greatly reduced in samples grown on Zn-exposed GaAs-c(4×4) with thermal annealing at 440 °C.

3. Origins of stacking faults from ZnSe/GaAs interface

As shown in Figs. 1 and 2 obtained from samples grown on Se- and Te-exposed GaAs, 3–4 monolayers of Ga₂Se₃- and Ga₂Te₃-like interfacial layers formed due to the interaction between Se or Te and the GaAs surfaces. This interaction is enhanced by thermal annealing of the Se- or Te-

exposed GaAs, so that the densities of defects increase (see Figs. 5(a) and 11). In addition, a high density of vacancy loops was generated due to thermal annealing in samples with Se- or Te-reacted interface. To form the 3–4 monolayers of Ga₂Se₃- or Ga₂Te₃-like interfacial layer, the formation of Ga vacancies may associate with the generation of Ga interstitials locally by a dissociative diffusion mechanism on the surfaces of the Se- or Te-exposed GaAs as is the case of heavily Se-doped GaAs.³² Since the lattice mismatch between ZnSe and GaAs is relatively small (~ 0.27%), the heterovalency of the ZnSe/GaAs interface may play a main driving force in the generation of the vacancies or interstitials. For instance, a Ga-Se or Ga-Te bonding may introduce 1/4 of an extra electron for the interaction between Se or Te atoms and the GaAs surface. Thus, formation of a Ga₂Se₃- or Ga₂Te₃-like interfacial layer is a relatively stable structure to balance the additional electrons due to Ga-Se or Ga-Te bondings. However, the interaction driven by the charge unbalance of Ga-Se or Ga-Te bondings may also associate with the formation of Ga vacancies and Ga interstitials due to the local misarrangement upon the formation of the Ga₂Se₃- or Ga₂Te₃-like interfacial layer. Consequently, intrinsic- and extrinsic-type stacking faults formed by following the same stacking sequence on these local stacking errors (see Fig. 5).

On the other hand, existence of ~ 2 and 3–4 monolayers of Zn-As interfacial layers were obtained in samples grown on Zn-exposed GaAs-(2×4) and c(4×4), respectively [see Figs. 3 and 4(a)]. In Zn-exposed GaAs-c(4×4), the top As layer on the GaAs-c(4×4) surface is probably replaced (or clustering) due to Zn exposure in order to form the 3–4 monolayers of Zn-As interfacial layer. In the case of samples grown on Zn-exposed As-terminated GaAs, the density of stacking faults and As precipitates increases as the interaction between Zn and excess As increases [see Fig. 9(a)]. During this interaction between Zn and excess As atoms, the Zn-As bondings may form with kicking the top surface of As by Zn atoms with the formation of local As interstitials or clustering on the GaAs surface such as the kick-out diffusion mechanism of Zn diffusion in GaAs.³³ However, the formation of the Zn-As bondings with a lack of electrons may provide a driving force for the generation of cation interstitials (such as Zn interstitials) for the charge balance on the GaAs surface. Therefore, clustering of excess As and formation of Zn interstitials on the GaAs surface takes place due to Zn exposure on the GaAs-c(4×4). These local clustering and stacking errors on the GaAs surfaces may act as nucleation sites for the generation of As precipitates and extrinsic-type stacking faults, respectively, after the growth of the ZnSe films. The interaction between Zn and excess As was decreased by thermal annealing with the decrease in As coverage on the GaAs surface and the thickness of the Zn-As interactive interfacial layers (see Fig. 4). In this case, a very low density of fault defects and a very smooth and very thin (close to intermixed) interfacial layer is obtained in samples grown on Zn-exposed GaAs-c(4×4) with 10–70 min thermal annealing at 440 °C [see Figs. 4(c)

and 12]. In this case, the interfacial layer has the chemical composition containing Ga, As, Zn, and Se, which may balance the interface charge and result in a decrease in the defect density. Therefore, the local stacking errors due to the interaction (or the charge unbalance) between Zn and excess As is reduced by thermal annealing and a very low density of the fault defects is obtained (see Fig. 12).

IV. CONCLUSIONS

We have investigated the quality of ZnSe/GaAs as a function of chemical and thermal preparation of GaAs surface to approach the origins for the generation of fault defects from the ZnSe/GaAs interface. Existence of 3–4 monolayers of Ga₂Se₃- and Ga₂Te₃-like interfacial layers is recognized under TEM 002- and 004-type reflection conditions in samples grown on Se- and Te-exposed GaAs surfaces with and without thermal annealing, respectively. Densities of extrinsic- and intrinsic-type stacking faults increased as the interaction between Se or Te atoms and GaAs surface increased by thermal annealing on the Se- or Te-exposed GaAs prior to the growth of the ZnSe films. In addition, a high density of vacancy loops of $\geq 5 \times 10^8/\text{cm}^2$ was generated in samples grown on Se- or Te-exposed GaAs with an annealing temperature of $\geq 400^\circ\text{C}$. Formation of intrinsic stacking faults or vacancy loops and extrinsic stacking faults may be related to the presence of cation vacancies and interstitials, respectively, on the surface of the GaAs epilayer by repeating the same stacking errors on these point defects during growth of the films. The interaction between the Se or Te and GaAs surface is associated with the formation of Ga vacancies and Ga interstitials due to the local misarrangement driven by charge unbalanced Ga–Se or Ga–Te bondings upon the formation of the Ga₂Se₃- or Ga₂Te₃-like interfacial layer. This interaction is enhanced by thermal annealing with an increase of charge unbalanced Ga–Se or Ga–Te bondings on the GaAs surface. Consequently, this interaction induces an increase in the nucleation sites (cation vacancies and interstitials) on the Se-exposed GaAs surface for the generation of high densities of vacancy loops, intrinsic-, and extrinsic-type stacking faults.

Existence of ~2 and 3–4 monolayers of Zn–As interfacial layers is suggested by TEM 002- and 004-type reflection images in samples grown on 1 min Zn exposure of As-rich GaAs-(2×4) and -c(4×4) surfaces, respectively. A very low density of stacking faults in the range of $10^4/\text{cm}^2$ was obtained in samples with Zn exposure on an As-stabilized GaAs-(2×4) surface. However, high densities of As precipitates and extrinsic-type stacking faults were generated from ZnSe/GaAs interface grown on As-rich GaAs-c(4×4) surfaces. The density of As precipitates increases as the surface As coverage of c(4×4) reconstruction increased on the As-stabilized GaAs-(2×4) surface and is associated with an increase in the density of extrinsic-type stacking faults bound by partial edge dislocations with a core structure terminated on additional cations. The formation of the 3–4 monolayers of the Zn–As interfacial layer may be associated with the

replacement or clustering of the top As due to Zn exposure on the GaAs-c(4×4). In this case, the formation of the Zn–As bondings with the lack of electrons may provide a driving force for the generation of cation interstitials (such as Zn interstitials) for the charge balance on the GaAs surface. Therefore, clustering of excess As atoms and formation of Zn interstitials on the GaAs surface may act as nucleation sites for the generation of the high densities of As precipitates and extrinsic-type stacking faults, respectively. The interaction between Zn and excess As on GaAs-c(4×4) is reduced by thermal annealing with a decrease in the thickness of the Zn–As interactive interfacial layer and desorption of excess As atoms from the GaAs surface. The local stacking errors generated due to this charge unbalanced interaction is reduced by thermal annealing and a very low density of the fault defects is obtained. These results indicate the importance on the desorption of excess As and prevention of Se contamination on an As-stabilized GaAs-(2×4) surface to obtain a high quality of ZnSe/GaAs.

ACKNOWLEDGMENTS

This work, partly supported by NEDO, was performed in the Joint Research Center for Atom Technology (JRCAT) under the joint research agreement between the National Institute for Advanced Interdisciplinary Research (NAIR) and the Angstrom Technology partnership (ATP).

- ¹P. M. Petroff and R. L. Hartman, *Appl. Phys. Lett.* **23**, 469 (1973).
- ²S. Mahajan, W. D. Johnston, Jr., M. A. Pollack, and R. E. Nahory, *Appl. Phys. Lett.* **34**, 717 (1979).
- ³S. Guha, J. M. DePuydt, M. A. Haase, J. Qiu, and H. Cheng, *Appl. Phys. Lett.* **63**, 3107 (1993).
- ⁴J. M. DePuydt, M. A. Haase, S. Guha, J. Qiu, H. Cheng, B. J. Wu, G. E. Hoffer, G. Meis-Haugen, M. S. Hagedorn, and P. F. Baude, *J. Cryst. Growth* **138**, 667 (1994).
- ⁵G. C. Hua, N. Otsuka, D. C. Grillo, Y. Fan, J. Han, M. D. Ringle, R. L. Gunshor, M. Hovinen, and A. V. Nurmikko, *Appl. Phys. Lett.* **65**, 1331 (1994).
- ⁶G. M. Haugen, S. Guha, H. Cheng, J. M. DePuydt, M. A. Haase, G. E. Hoffer, J. Qiu, and B. J. Wu, *Appl. Phys. Lett.* **66**, 358 (1995).
- ⁷S. Tomiya, E. Morita, M. Ukita, H. Okuyama, S. Itoh, K. Nakano, and A. Ishibashi, *Appl. Phys. Lett.* **66**, 1208 (1995).
- ⁸L. H. Kuo, L. Salamanca-Riba, B. J. Wu, G. M. Haugen, J. M. DePuydt, G. Hoffer, and H. Cheng, *J. Vac. Sci. Technol. B* **13**, 1694 (1995).
- ⁹D. J. Eaglesham, D. P. Kvam, D. M. Bean, J. C. Bean, and C. J. Humphreys, *Philos. Mag.* **59**, 1059 (1989).
- ¹⁰L. H. Kuo, L. Salamanca-Riba, J. M. DePuydt, H. Cheng, and J. Qiu, *Philos. Mag.* **69**, 301 (1994).
- ¹¹L. H. Kuo, L. Salamanca-Riba, G. Hoffer, and B. J. Wu, *Philos. Mag. A* **71**, 883 (1995).
- ¹²D. Cherns, S. Mylonas, D. E. Ashenford, and B. Lunn, *Philos. Mag. A* **74**, 553 (1996).
- ¹³E. A. Fitzgerald, G. P. Watson, R. E. Proano, D. G. Ast, P. D. Kirchner, G. D. Pettit, and J. M. Wooddall, *J. Appl. Phys.* **65**, 2220 (1989).
- ¹⁴L. H. Kuo, L. Salamanca-Riba, J. M. DePuydt, H. Cheng, and J. Qiu, *Appl. Phys. Lett.* **63**, 3197 (1993).
- ¹⁵L. H. Kuo, L. Salamanca-Riba, J. M. DePuydt, H. Cheng, and J. Qiu, *J. Electron. Mater.* **23**, 275 (1994).
- ¹⁶L. H. Kuo, L. Salamanca-Riba, B. J. Wu, G. Hoffer, J. M. DePuydt, and H. Cheng, *Appl. Phys. Lett.* **67**, 3298 (1995).
- ¹⁷L. H. Kuo, K. Kimura, T. Yasuda, S. Miwa, C. G. Jin, K. Tanaka, and T. Yao, *Appl. Phys. Lett.* **68**, 2413 (1996).
- ¹⁸B. J. Wu, G. M. Haugen, J. M. DePuydt, L. H. Kuo, and L. Salamanca-Riba, *Appl. Phys. Lett.* **68**, 2828 (1996).

- ¹⁹L. H. Kuo, K. Kimura, S. Miwa, T. Yasuda, and T. Yao, *J. Electron. Mater.* **26**, 52 (1997).
- ²⁰Edith D. Bourret-Courchesne, *Appl. Phys. Lett.* **68**, 1675 (1996).
- ²¹L. H. Kuo, K. Kimura, S. Miwa, T. Yasuda, and T. Yao, *Appl. Phys. Lett.* **69**, 1408 (1996).
- ²²J. P. Hirth and J. Lothe, *Theory of Dislocations*, 2nd ed. (Wiley, New York, 1982), pp. 373–381.
- ²³P. Haasen, *Acta Metall.* **5**, 598 (1957).
- ²⁴T. Yasuda, L. H. Kuo, K. Kimura, S. Miwa, C. Jin, K. Tanaka, and T. Yao, *J. Vac. Sci. Technol. B* **14**, 3052 (1996).
- ²⁵D. K. Biegelsen, R. D. Bringans, J. E. Northrup, and L. E. Swartz, *Phys. Rev.* **41**, 5701 (1990).
- ²⁶A. Ohtake, L. H. Kuo, T. Yasuda, K. Kimura, S. Miwa, and T. Yao, *J. Vac. Sci. Technol.* (submitted).
- ²⁷D. Li, J. M. Gonsalves, N. Otsuka, J. Qiu, M. Kobayashi, and R. L. Gunshor, *Appl. Phys. Lett.* **57**, 449 (1990).
- ²⁸*International Tables for X-ray Crystallography* (Kynoch Press, Birmingham, England, 1962), Vol. III, p. 210.
- ²⁹S. Heun, J. J. Paggel, S. Rubini, and A. Franciosi, *J. Vac. Sci. Technol. B* **14**, 2980 (1996).
- ³⁰S. P. Hirsch, A. Howie, R. B. Nicholson, D. W. Pashley, and M. J. Whelan, *Electron Microscopy of Thin Crystals* (Krieger, New York, 1977), p. 327.
- ³¹R. Hull, J. M. Gibson, and J. C. Bean, *Appl. Phys. Lett.* **46**, 179 (1985).
- ³²M. S. Abrahams, J. Blanc, and C. J. Buiochi, *J. Appl. Phys.* **45**, 3277 (1974).
- ³³A. H. van Ommen, *J. Appl. Phys.* **54**, 5055 (1983).

Growth mode and defect generation in ZnSe heteroepitaxy on Te-terminated GaAs(001) surfaces

A. Ohtake,^{a)} L. H. Kuo,^{a)} T. Yasuda,^{b)} K. Kimura,^{a)} S. Miwa,^{a)} and T. Yao^{b),c)}
Joint Research Center for Atom Technology (JRCAT), Tsukuba 305, Japan

K. Nakajima and K. Kimura

Department of Engineering Physics and Mechanics, Kyoto University, Kyoto 606-01, Japan

(Received 12 January 1997; accepted 10 April 1997)

We have studied growth mode and defect generation in heteroepitaxy of ZnSe on Te-terminated GaAs(001) surfaces. The high saturation coverage of Zn on Te-terminated GaAs enhances the layer-by-layer growth of ZnSe. However, high densities of faulted defects ($\sim 5 \times 10^8/\text{cm}^2$) are generated in the ZnSe film. We have found that the generation of defects is not necessarily ascribed to an island growth mode, but is closely related to the formation of a vacancy-contained Ga-Te interface layer. © 1997 American Vacuum Society. [S0734-211X(97)07604-X]

I. INTRODUCTION

Heterojunctions between different semiconductor materials hold significant potential for technological applications in devices. It is sufficiently understood that performance and reliability of a device strongly depend on the structural quality of heterojunctions. A primary concern in obtaining high-quality heterojunctions is the growth mode of epitaxial layers. Epitaxial growth has led to general classification into three growth modes; layer-by-layer (Frank-van der Merwe), islanding (Volmer-Wever), or layer-by-layer followed by islanding (Stranski-Krastanov). Since the coalescence of islands is generally thought to lead to the generation of defects, the layer-by-layer growth mode has been desired.

For heteroepitaxy of ZnSe on GaAs(001), there have been numerous studies showing that the density and type of defects in the film strongly depend on the initial growth mode.¹⁻³ Among them, it is interesting to point out that Se and Te, both being the column VI elements, have the opposite effects on ZnSe growth. Island growth of ZnSe is caused by an initial exposure of the GaAs surface to a Se beam prior to the growth, giving rise to the generation of high densities of defects.^{1,2} On the other hand, Spahn *et al.* recently reported that the termination of the GaAs surface with Te enhances layer-by-layer growth of ZnSe, which, they claimed, improves the crystallinity of the ZnSe film.³

The purpose of this article is (i) to study the growth mode of ZnSe on the Te-terminated GaAs surface, and (ii) to discuss the correlation between the growth mode and defect generation in the ZnSe film. Layer-by-layer growth of ZnSe was confirmed by intensity oscillations in reflection high-energy electron diffraction (RHEED). We have clarified that the majority of Te atoms remain at the ZnSe/GaAs interface using high-resolution Rutherford backscattering spectroscopy (RBS) and reflectance difference spectroscopy (RDS).

Transmission electron microscopy (TEM) observations have shown that high densities of faulted defects are generated in the ZnSe film grown on the Te-terminated surface, even when the growth proceeds in the layer-by-layer mode.

II. EXPERIMENT

The experiments were performed in a dual-chamber molecular-beam epitaxy (MBE) system equipped with *in situ* RDS, RHEED, x-ray photoelectron spectroscopy (XPS), and scanning tunneling microscopy apparatuses. A detailed description of the experimental apparatuses used in this study has been given elsewhere.^{4,5} After growth of an undoped homoepitaxial layer ($\sim 0.5 \mu\text{m}$) on a thermally cleaned GaAs(001) substrate, the surface showed an As-stabilized (2×4) reconstruction, as confirmed by RHEED observations and RDS measurements.⁴ The sample was then transferred via ultrahigh vacuum transfer modules to another MBE chamber for a Te exposure followed by ZnSe growth. The sample was exposed to a Te beam with a beam-equivalent pressure (BEP) of 1.1×10^{-7} Torr for 60 s at 400 °C. The coverage of Te is estimated from the area of the Te $3p_{1/2}$ and $3p_{3/2}$ XPS peaks, after calibration using *ex situ* RBS measurements. Thin pseudomorphic ZnSe films ($< 1000 \text{ \AA}$) were grown on Te-terminated GaAs surfaces at 250 °C. The BEPs for Zn and Se were 4.0×10^{-7} and 9.0×10^{-7} , respectively. The growth rate of ZnSe was 1.1 \AA/s under these conditions. For comparison, we performed ZnSe growth experiments on the Se-terminated surfaces prepared by exposure of Se on the GaAs surface for 60 s at 250 °C with and without subsequent thermal annealing at 520 °C.

In RDS, the directly measured quantity is anisotropy in the complex reflectance $\Delta \tilde{r}/r = (\tilde{r}_{1\bar{1}0} - \tilde{r}_{110})/\tilde{r} = \Delta r/r + i\Delta\theta$, where $\tilde{r}_{1\bar{1}0}$ and \tilde{r}_{110} are the near-normal-incidence complex reflectances for light linearly polarized along $[1\bar{1}0]$ and $[110]$, respectively. Here, we present only the data in the form $\Delta r/r = \text{Re}(\Delta \tilde{r}/\tilde{r})$. RHEED patterns were taken along the $\langle 110 \rangle$ and $\langle 100 \rangle$ azimuth of GaAs(001) with an electron-beam energy of 10 keV. XPS measurements were

^{a)}Also with Angstrom Technology Partnership (ATP), Tsukuba 305, Japan; Electronic mail: take@jrcat.or.jp

^{b)}Also with National Institute for Advanced Interdisciplinary Research (NAIR), Tsukuba 305, Japan.

^{c)}Also with Institute for Materials Research, Tohoku University, Sendai 980, Japan.

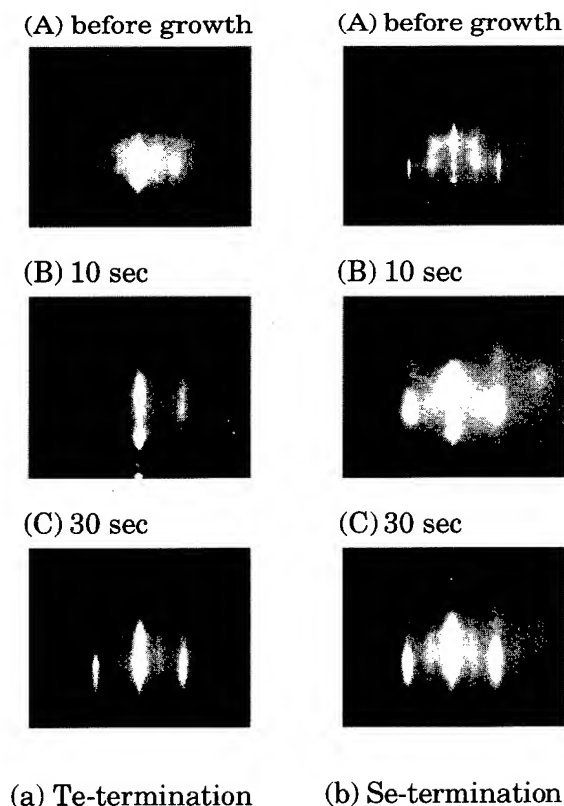


FIG. 1. A series of RHEED patterns taken along [110] azimuth of ZnSe films growing on the Te-terminated (a) and the Se-terminated (b) GaAs(001) surfaces, respectively.

carried out by using monochromatic Al $K\alpha$ radiation (1486.6 eV). The samples were also characterized by TEM to assess the density and type of defects in ZnSe films and the interface layer of ZnSe/GaAs, and by high-resolution RBS to measure depth profiles of Te in ZnSe films. The high-resolution RBS measurements were performed with the system at Kyoto University.⁶

III. RESULTS AND DISCUSSION

A. Growth mode of ZnSe on Te-terminated GaAs

When the GaAs(001)-(2 \times 4) surface was exposed to the Te beam, the RHEED pattern changed to one exhibiting a (6 \times 1) reconstruction [Fig. 1(a)-A]. On the basis of XPS and RBS measurements, the Te coverage of the (6 \times 1) surface is estimated to be ~ 1.2 ML (1 ML = 6.2×10^{14} atom/cm²), in agreement with previously reported results.⁷

Figures 1(a) and 1(b) show a series of RHEED patterns taken along the [110] azimuth of the growing ZnSe films on the Te- and Se-terminated (without annealing) GaAs(001) surfaces, respectively. ZnSe grown on Te-terminated GaAs shows streaked RHEED patterns from the beginning of the growth [Fig. 1(a)], while those from ZnSe on Se-terminated GaAs consist of spots [Fig. 1(b)]. Figure 2 shows a RHEED intensity oscillation profile observed while ZnSe is growing on Te-terminated GaAs. The glancing angle of the incident electron beam was approximately 1.6° (so-called off-Bragg

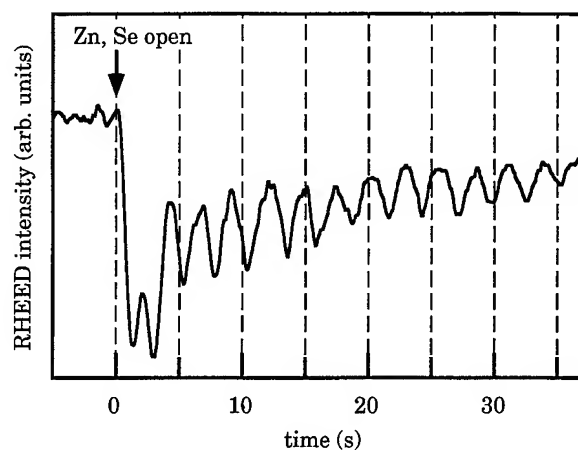


FIG. 2. RHEED intensity oscillations obtained along the [110] direction while ZnSe grows on Te-terminated GaAs.

condition). Under such a condition, RHEED intensity is sensitive to surface roughness and growth modes. In Fig. 2, oscillations with a period corresponding to the bilayer growth of ZnSe(001) are observed. These results confirm that the Te- and Se-terminations of the initial surfaces result in a layer-by-layer growth mode and an island growth mode, respectively.

Such an effect of the Te atoms on the growth mode of ZnSe reminds us what is known for so-called surfactant-mediated epitaxy. In surfactant-mediated epitaxy, an artificially deposited surfactant material (e.g., As in Ge/Si) segregates to the growth front, promoting layer-by-layer growth.⁸ In order to address the question as to whether Te segregates to the growth front to serve as a surfactant in this system, RBS measurements were performed. Figure 3 shows the high-resolution RB spectrum obtained for the ZnSe film (~ 90 Å thick) grown on Te-terminated GaAs. The spectrum shows a well-defined peak at 315 keV, which corresponds to

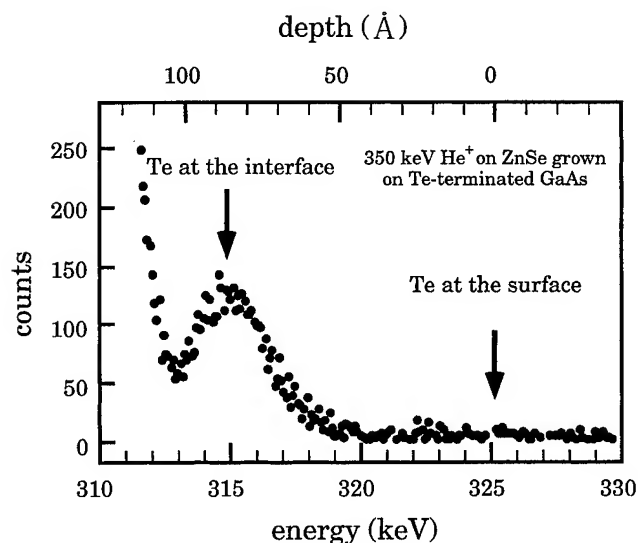


FIG. 3. High-resolution RB spectrum obtained for the ZnSe film (90 Å thick) grown on the Te-terminated GaAs(001) surface.

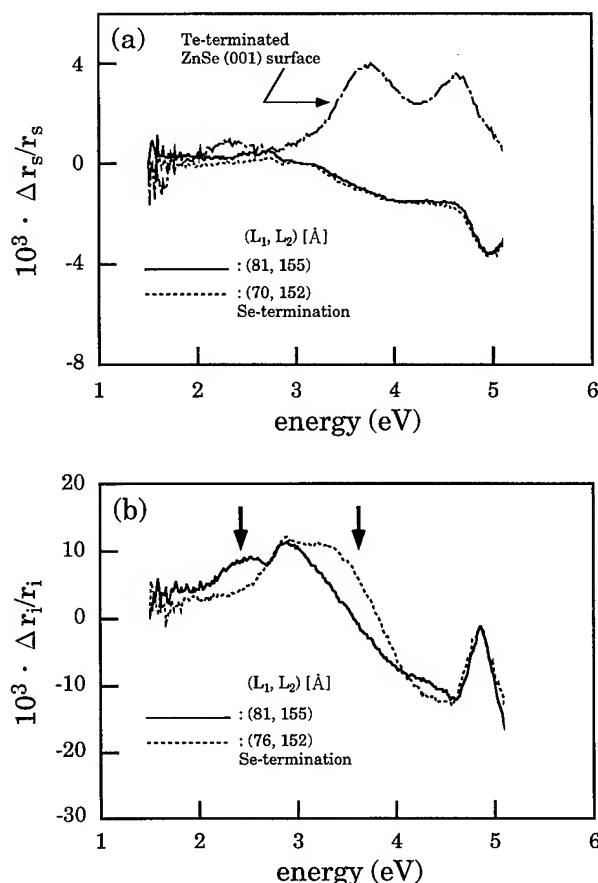


FIG. 4. The surface (a) and interface (b) contributions to the measured RD spectra. Also plotted in (a) and (b) are the results for growth on the Se-terminated surface (dotted curve). The spectra were obtained by using two RDS measurements at different film thicknesses, L_1 and L_2 , labeled in the figure legend. Since the exposure of the Te-terminated surface to the Se beam alone can eliminate surface Te atoms (see the text), the RDS measurements were performed only under the Zn beam.

the ions scattered from Te atoms remaining at the ZnSe/GaAs interface. On the other hand, we could not obtain convincing evidence for the existence of Te at the ZnSe surface (~ 325 keV) and/or in the ZnSe film. The amount of Te at the interface is estimated to be ~ 1.1 ML, which agrees with the Te coverage on the GaAs(001)-(6 \times 1)-Te surface (~ 1.2 ML). This means that the majority of Te atoms on the initial surface remain at the interface even after the ZnSe growth. Details of the RBS studies will be published elsewhere.⁹

RDS measurements also confirmed that Te remains near the interface. Figures 4(a) and 4(b) show the RD spectra of the ZnSe surface and the ZnSe/GaAs interface, respectively, which were obtained by using the procedure developed by Yasuda *et al.*⁴ In Figs. 4(a) and 4(b), the results for growth on Se-terminated GaAs (dotted curve) are also shown for comparison. The RD spectrum of the interface, $\Delta r_i/r_i$ [Fig. 4(b)] are significantly altered by Te, as indicated by arrows in Fig. 4(b). On the other hand, the surface RD spectrum, $\Delta r_s/r_s$, [Fig. 4(a)] agrees well with that for Se termination. We have confirmed that the Te termination of the ZnSe(001) surface caused a dramatic change in the line shape of the

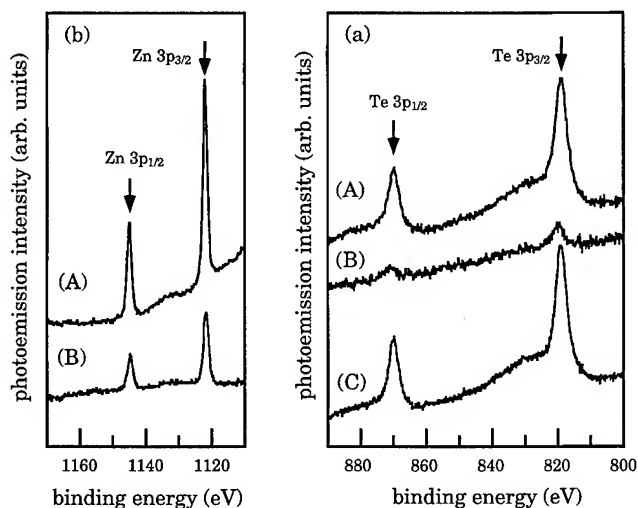


FIG. 5. (a) Te $3p_{1/2}$ and $3p_{3/2}$ photoemission spectra measured before (A) and after the Te-terminated GaAs(001) surface was exposed to the Se (B) and Zn (C) beams. (b) Zn $3p_{1/2}$ and $3p_{3/2}$ photoemission spectra measured after Zn exposure on the Te-terminated GaAs(001) (A) and GaAs(001)-(2 \times 4) (B) surfaces.

spectrum, as shown by the dashed-dotted curve in Fig. 4(a). Thus, the possibility of Te segregation at the ZnSe surface during the growth can be dismissed. From RBS and RDS results, we conclude that Te does not segregate to the growth front, but remains near the ZnSe/GaAs interface.

B. Role of Te in the layer-by-layer growth of ZnSe

Since Te does not segregate to the growth front, the role of Te in enhancing layer-by-layer growth is distinctly different from that for surfactant-mediated epitaxy; for example, InAs on Te-terminated GaAs(001) (Ref. 10) and Ge on Te-terminated Si(001).¹¹ In order to clarify the role of Te at the initial stage of ZnSe growth, we performed ZnSe growth experiments after the Te-terminated surface was exposed to either the Se or the Zn beam at 250 °C. Figures 5(a)-A and 5(a)-B show photoemission spectra measured before and after the Se exposure on the Te-terminated GaAs surface, respectively. As is evident from Figs. 5(a)-A and 5(a)-B, intensities of the Te $3p_{1/2}$ and $3p_{3/2}$ peaks are decreased by Se exposure. This means that the majority of Te atoms are replaced by Se. Indeed, our RBS measurements showed that ~ 0.8 ML of Te was desorbed after Se exposure. This was also evidenced by the RDS measurement. Shown in a part of Fig. 6 is the RD spectrum for the Te-terminated surface (solid curve). Characteristic features for the Te-terminated surface disappeared upon Se exposure: the spectrum measured after Se exposure (dashed-dotted curve) agrees well with that for the Se-terminated surface (dotted curve) in both line shape and intensity. In addition, the RHEED observation showed that the Se exposure transformed the surface reconstruction from the (6 \times 1) to (2 \times 1) characteristic of the Se-terminated GaAs(001) surface.^{4,12} These results are consis-

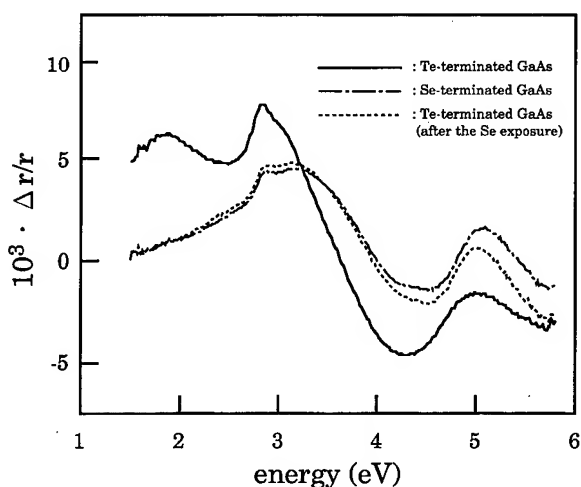


FIG. 6. RD spectra for the Te- and Se-terminated GaAs(001) surfaces. Also shown in the figure is the RD spectrum measured after the Te-terminated surface was exposed to the Se beam.

tent with the result of first-principles pseudopotential calculations: Se is bonded stronger than Te to the GaAs surface.¹³

As predicted from these results, our RHEED observations showed that the ZnSe growth initiated with Se exposure on Te-terminated GaAs proceeds in island formation, similar to the growth on Se-terminated GaAs [Fig. 1(b)]. In addition, our RBS measurements showed that only ~ 0.3 ML of Te remains at the interface. These results stand in clear contrast to the result described in the preceding subsection, where the growth is initiated by simultaneous exposure of both Zn and Se beams: ZnSe grows in the layer-by-layer mode, and the majority of the Te atoms (~ 1.1 ML) remains at the interface. On the other hand, when the Te-terminated surface was exposed to the Zn beam, as shown in Fig. 5(a)-C, the $3p_{1/2}$ and $3p_{3/2}$ intensities remained almost unchanged. Also, we have confirmed that the subsequent ZnSe growth proceeds in the layer-by-layer mode, leaving ~ 1.2 ML of Te at the interface. From these results, it appears likely that the layer-by-layer growth of ZnSe on the Te-terminated GaAs(001) surface begins with the adsorption of Zn.

In addition, comparing carefully the Zn $3p_{1/2}$ and $3p_{3/2}$ intensities measured after Zn exposure on the Te-terminated GaAs surface [Fig. 5(b)-A] with those for the Zn-treated GaAs(001)-(2 \times 4) surface [Fig. 5(b)-B], we found that the Te termination doubles the saturation coverage of Zn. On the other hand, our XPS measurements confirmed that Zn atoms are hardly adsorbed on the Se-terminated GaAs(001)-(2 \times 1) surface. Thus, we speculate that the high saturation coverage of Zn on Te-terminated GaAs increases the density of the two-dimensional nuclei of ZnSe, leading to layer-by-layer growth.

C. Defect generation in ZnSe grown on Te-terminated GaAs

Figure 7(a) shows a typical plan-view TEM image of the ZnSe film (~ 1000 Å thick) grown on the Te-terminated

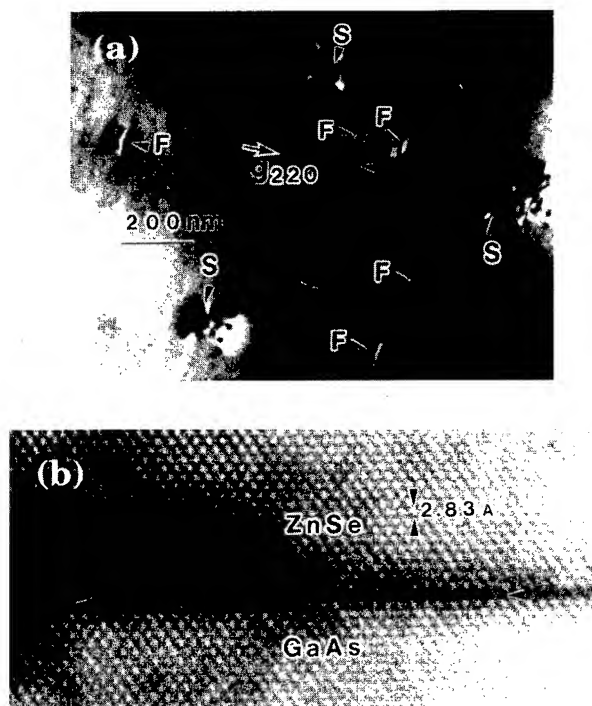


FIG. 7. (a) TEM plan-view 220 weak beam image of the ZnSe film grown on the Te-terminated GaAs(001) surface. (b) High-resolution TEM image of the ZnSe/GaAs interface formed on the Te-terminated GaAs surface.

GaAs(001) surface. High densities of intrinsic Frank- and extrinsic Shockley-type partial dislocations [labeled F and S in Fig. 7(a), respectively] are observed. The densities of these defects are both $\sim 5 \times 10^8/\text{cm}^2$. In addition, a very high density of Frank-type vacancy loops, marked by arrowheads in Fig. 7(a), is observed ($> 1 \times 10^9/\text{cm}^2$). As shown in Table I, these densities are slightly higher than those in ZnSe films grown on the Se-terminated GaAs surface,¹⁴ and are much higher than those ($< 1 \times 10^4/\text{cm}^2$) obtained by the initial Zn exposure on the GaAs(001)-(2 \times 4) surface.^{2,14,15} This means that the Te termination of the GaAs(001) surface is useless in suppressing the defect generation.

The generation of high densities of defects in the ZnSe film grown on Te-terminated GaAs seems to disagree with the results of Spahn *et al.*:³ they claimed that the Te termination improves the crystalline quality of the ZnSe film on the basis of their high-resolution x-ray diffraction (XRD) measurements. However, they did not measure defect densities. Although the degradation in crystallinity caused by an introduction of dislocations usually results in some peak broadening in the measured XRD profiles, XRD is less sensitive to the existence of microscopic defects, such as stacking faults. Considering that their XRD measurements were performed for pseudomorphic (dislocation-free) ZnSe films, it is possible that samples containing high densities of defects exhibit XRD profile with a narrow width.

As mentioned in the introduction, islanding, which frequently occurs in heteroepitaxy, is often an undesirable growth mode, because it tends to degrade the crystalline quality of growing films. Thus, the ZnSe film grown in a

TABLE I. Density of faulted defects in ZnSe grown on the Te- and Se-terminated GaAs surfaces.

	Te termination	Se termination (without annealing) ^a	Se termination (with annealing) ^a
Frank-type partial dislocations (/cm ²)	5×10 ⁸	5×10 ⁷	1×10 ⁸
Shockley-type partial dislocations (/cm ²)	5×10 ⁸	5×10 ⁷	5×10 ⁸
Frank-type vacancy loops (/cm ²)	>1×10 ⁹	<5×10 ⁴	1×10 ⁹

^aSee Ref. 14.

layer-by-layer mode on Te-terminated GaAs might be expected to contain a low density of defects when compared to the case of Se termination (island growth). However, the present results clearly show that the generation of defects is not necessarily ascribed to the island growth mode.

Shown in Fig. 7(b) is a high-resolution lattice image of the {110} cross-sectional sample, in which a layer with a width of about 4 ML is seen along the ZnSe/GaAs interface. Lattice fringes in Fig. 7(b) show a perfect coherency between the ZnSe and GaAs crystal lattices, indicating that a zinc-blende structure with a different structure factor is formed at the interface. Our scattering factor analysis on 002- and 004-type dark field images suggests that this layer is a Ga-Te compound containing cation vacancies.⁹ A known example of such a compound is Ga₂Te₃.

From our TEM scattering factor analysis, no evidence of the existence of a Zn-Te interface layer can be suggested.^{9,16} This seems reasonable by considering that the heat of the formation of Ga₂Te₃ and ZnTe are -65.7 kcal/mol (Ref. 17) and -24.9 kcal/mol,¹⁸ respectively. Here, it should be noted that our XPS results showed that ZnSe growth is initiated with the adsorption of Zn on the Te-terminated surface, as mentioned earlier. This means that the interface between ZnSe and Te-terminated GaAs is abrupt, namely, the uptake of Te in ZnSe is negligible. On the other hand, since the width of the Ga-Te layer is about 4 ML, it is likely that Te undergoes anion exchange down to a depth of ~4 ML into the GaAs substrate. This is in contrast to the work reported by Chambers and Sundaram,¹⁹ where the indiffusion of Te into subsurface layers of GaAs(001) was very limited. The reason for this discrepancy may be due to the different surface treatment: Chambers and Sundaram treated the GaAs surface with (C₂H₅)₂Te by metal-organic chemical vapor deposition.

It is known that a vacancy-contained Ga₂Se₃-like layer is formed at the interface on Se-terminated GaAs.²⁰ Kuo and co-workers reported that such an interface layer is the source for the formation of stacking faults because it generates point defects such as vacancies, which locally form stacking errors at the interface.^{14,15} Cherns *et al.*²¹ recently reported that in the ZnTe/GaSb system vacancy-type defects arise owing to a supersaturation of vacancies generated in an interface Ga₂Te₃ phase. It is likely, therefore, that the generation of faulted defects in this system is closely associated with the formation of the vacancy-contained Ga-Te interface layer. In addition, we point out that the covalent radius of Te (1.32

Å) is larger than those of As (1.18 Å) and Se (1.14 Å).²² Thus, the interface layers formed on GaAs would have larger stress for Te than Se, which may further induce the defect generation.

IV. CONCLUSION

We have systematically studied growth mode and defect generation in ZnSe heteroepitaxy on Te-terminated GaAs(001) surfaces. The ZnSe growth proceeds in the layer-by-layer mode, which is ascribed to the high saturation coverage of Zn on the Te-terminated GaAs surface. During the ZnSe growth, Te does not segregate to the growth front, but remains at the ZnSe/GaAs interface. The Te atoms at the interface form a vacancy-contained Ga-Te interface layer, which becomes the source for the generation of high densities of faulted defects in the ZnSe film.

ACKNOWLEDGMENTS

This study, partly supported by the New Industrial Technology Development Organization (NEDO), was performed at the Joint Research Center for Atom Technology (JRCAT) under the research agreement between the National Institute for Advanced Interdisciplinary Research (NAIR) and the Angstrom Technology Partnership (ATP). The authors would like to acknowledge Dr. H. Kobayashi of the SONY corporation and Professor F. Nishiyama of Hiroshima University for the RBS measurements.

¹S. Guha, H. Muneke, F. K. Legoues, and L. L. Chang, Appl. Phys. Lett. **60**, 3220 (1992); S. Guha, H. Muneke, and L. L. Chang, J. Appl. Phys. **73**, 2294 (1993).

²J. M. Gaines, J. Petruzzello, and B. Greenberg, J. Appl. Phys. **73**, 2835 (1993).

³W. Spahn, H. R. Ress, K. Schüll, M. Ehinger, D. Hommel, and G. Landwehr, J. Cryst. Growth **159**, 761 (1996).

⁴T. Yasuda, K. Kimura, S. Miwa, L. H. Kuo, C. G. Jin, K. Tanaka, and T. Yao, Phys. Rev. Lett. **77**, 326 (1996); J. Vac. Sci. Technol. **B 14**, 3052 (1996).

⁵S. Miwa, L. H. Kuo, K. Kimura, T. Yasuda, and T. Yao, Appl. Surf. Sci. **36**, L337 (1997).

⁶K. Kimura, K. Ohshima, and M. Mannami, Appl. Phys. Lett. **64**, 2232 (1994).

⁷Y. Gobil, J. Cibert, K. Saminadayar, and S. Tatarenko, Surf. Sci. **211/212**, 969 (1989); V. H. Etgens, R. Pinchaux, M. Sauvage-Simkin, J. Massies, N. Jedrecy, N. Greiser, and S. Tatarenko, *ibid.* **251/252**, 478 (1991).

⁸M. Copel, M. C. Reuter, E. Kaxiras, and R. M. Tromp, Phys. Rev. Lett. **63**, 632 (1989).

⁹A. Ohtake, L. H. Kuo, T. Yasuda, K. Kimura, S. Miwa, C. G. Jin, K. Nakajima, K. Kimura, and T. Yao (unpublished).

- ¹⁰N. Grandjean, J. Massies, and V. H. Etgens, *Phys. Rev. Lett.* **69**, 796 (1992); J. Massies, N. Grandjean, and V. H. Etgens, *Appl. Phys. Lett.* **61**, 99 (1992); J. Massies and N. Grandjean, *Phys. Rev. B* **48**, 8502 (1993).
- ¹¹S. Higuchi and Y. Nakanishi, *J. Appl. Phys.* **71**, 4277 (1992).
- ¹²D. Li and M. D. Pashley, *J. Vac. Sci. Technol. B* **12**, 2547 (1994).
- ¹³T. Ohno, *Surf. Sci.* **255**, 229 (1991).
- ¹⁴L. H. Kuo, K. Kimura, T. Yasuda, S. Miwa, C. G. Jin, K. Tanaka, and T. Yao, *Appl. Phys. Lett.* **68**, 2413 (1996).
- ¹⁵L. H. Kuo, L. Salamanca-Riba, B. J. Wu, G. Hofler, J. M. Depuydt, and H. Cheng, *Appl. Phys. Lett.* **67**, 3298 (1995); L. H. Kuo, L. Salamanca-Riba, B. J. Wu, G. M. Haugen, J. M. Depuydt, G. Hofler, and H. Cheng, *J. Vac. Sci. Technol. B* **13**, 1694 (1995).
- ¹⁶The 004 structure factor of Ga₂Te₃ is almost the same as those of ZnSe and GaAs, which is consistent with the present TEM observation. On the other hand, the 004 structure factor for ZnTe is larger than those of ZnSe and GaAs.
- ¹⁷A. S. Abbasov, A. V. Nikolskaya, Y. I. Gerasimov, and V. P. Vasilev, *Dokl. Akad. Nauk. SSSR* **156**, 1140 (1964).
- ¹⁸P. Goldfinger and M. Jeunhomme, *Trans. Faraday Soc.* **59**, 2851 (1963).
- ¹⁹S. A. Chambers and V. S. Sundaram, *J. Vac. Sci. Technol. B* **9**, 2256 (1991).
- ²⁰D. Li, J. M. Gonsalves, N. Otsuka, J. Qiu, M. Kobayashi, and R. L. Gunshor, *Appl. Phys. Lett.* **57**, 449 (1990).
- ²¹D. Cherns, S. Mylonas, D. E. Ashenford, and B. Lunn, *Philos. Mag. A* **74**, 553 (1996).
- ²²C. Kittel, *Introduction to Solid State Physics*, 6th ed. (Wiley, New York, 1986).

In situ* photoemission and reflectance anisotropy spectroscopy studies of CdS grown on InP(001)

Ch. Schultz, A. M. Frisch, K. Hinrichs, J. Kinsky, T. Herrmann, U. Rossow,^{a)} N. Esser, and W. Richter
Technische Universität Berlin, Institut für Festkörperphysik, Hardenbergstrasse 36, D-10623 Berlin, Germany

(Received 13 January 1997; accepted 5 May 1997)

Structural and electronic properties of the clean InP(001) surface and the CdS/InP(001) interface were investigated by soft x-ray photoemission and reflectance anisotropy spectroscopy (RAS). Clean InP(001) (2×4) surfaces were prepared by desorption of a protective arsenic/phosphorus double capping layer, thereafter, CdS deposited by single source molecular beam epitaxy at a substrate temperature of 200 °C. Core-level and valence-band photoemission spectra, as well as the RAS data, were taken *in situ*. Core-level spectra show a disappearance of In 4*d* and P 2*p* surface core-level shifts and a chemical reaction between indium and sulphur upon CdS deposition. From valence-band spectra, a total valence-band offset of $\Delta E_V = (0.75 \pm 0.10)$ eV is derived, which corresponds to previous results of the CdS/InP(110) interface. After deposition of thin CdS layers, RAS spectra show strong features close to interband critical points of bulk InP due to surface roughening induced by the interface reaction. A feature around 5.3 eV, developing with higher coverage, indicates the formation of the metastable cubic phase of CdS as confirmed by additional Raman spectra. © 1997 American Vacuum Society. [S0734-211X(97)11504-9]

I. INTRODUCTION

For technological applications of semiconductor heterostructures, the knowledge of interface properties is of great importance. The combination of photoemission studies on the one hand and surface sensitive optical methods on the other hand is a very good tool to investigate the structural, chemical, and electronic properties of semiconductor heterojunctions. The band discontinuities of heterojunctions are hereby key design parameters, since the valence- and conduction-band offsets, ΔE_V and ΔE_C , determine the transport and confinement properties at the interface.

Recently, the CdS/InP(110) heterojunction has been well characterized. An interfacial reaction between indium and sulfur was observed by photoelectron spectroscopy and a valence-band offset of -0.77 ± 0.1 eV was derived.¹ CdS grows on the (110) surface of InP in the metastable cubic modification up to a layer thickness in the nanometer range.^{2,3} Investigations of the growth of CdS on the more technologically relevant clean InP(001) surface, however, are still missing. To understand the electronic properties of heterostructures, it is of fundamental interest to determine the band offsets for different crystal orientations to investigate whether commutativity and transitivity rules for wide band-gap heterojunctions are valid.⁴ Comparison of (110) and (100) interfaces would yield information regarding the effects of the orientation or interface reactions, which are neglected in these rules.

II. EXPERIMENT

The experiments were carried out with an angle-resolved photoelectron spectrometer (VG ADES 400) at the TGM 2-beamline of the Berliner Elektronen-Speicherring-Gesellschaft für Synchrotronstrahlung storage ring in Berlin. All photoemission spectra were taken in normal emission with an overall resolution <0.2 eV at a photon energy of 55 eV. A gold foil mounted in electrical contact with the sample was measured as an energy reference. In order to determine the energies of the core-level lines and the valence-band maximum with respect to the Fermi energy E_F , spectra were recorded from a gold foil mounted in electrical contact with the sample. The ultrahigh vacuum chamber (base pressure: 1×10^{-8} Pa) was also equipped with a low-energy electron diffraction (LEED) optics and a strain-free window for the reflectance anisotropy spectroscopy (RAS) measurements. The setup of the RAS is described in detail elsewhere.⁵

The InP(001) surfaces were prepared by thermal desorption of a protective arsenic/phosphorus sandwich layer. This procedure was controlled by quadrupole mass spectroscopy and the pressure of the chamber. To avoid contaminations on the clean surface due to the annealing procedure, a molybdenum sample holder was used. First, the arsenic layer is desorbed at a temperature of 550 K, at 590 K the phosphorous layer desorbs. CdS layers were deposited onto the clean InP(001) surface from CdS powder in a single Knudsen cell at a crucible temperature of 720 K. The sample temperature was 470 K.

III. RESULTS AND DISCUSSION

A. Growth of CdS on InP(001)

After desorption of the protective arsenic/phosphorous sandwich layer, the clean InP(001) surface forms a well-ordered (2×4) reconstruction as observed by LEED. The

*No proof corrections received from author prior to publication.

^{a)}Present address: TU Ilmenau, Inst. f. Physik, Pobox 100565, D-98684 Ilmenau, Germany; Electronic mail: frosch@mail.physik.tu-berlin.de

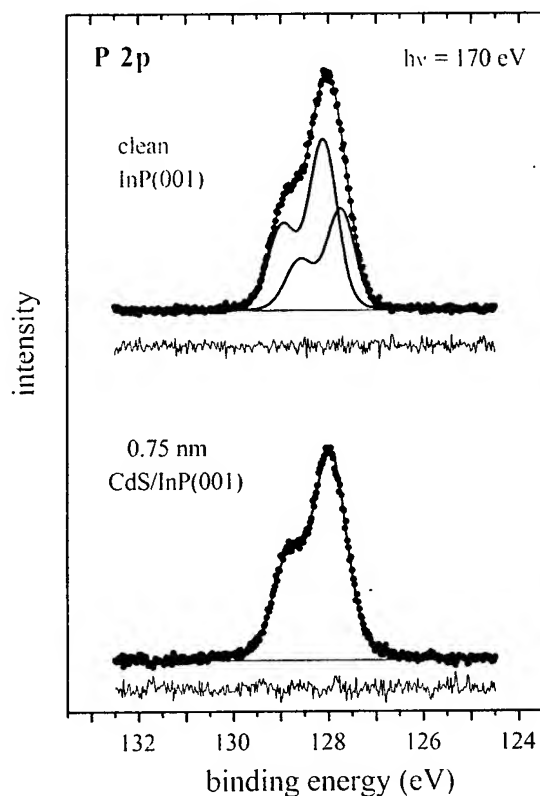


FIG. 1. Photoelectron spectra of the P $2p$ level of the clean and CdS covered InP(001) surface, recorded with a photon energy of 170 eV. For both spectra, the residual is shown at the bottom.

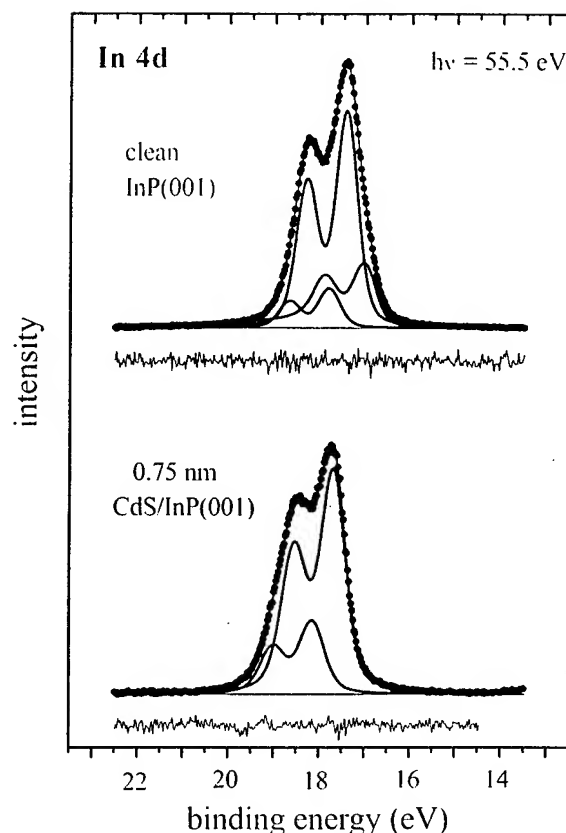


FIG. 2. Photoelectron spectra of the In $4d$ level of the clean and CdS covered InP(001) surface, recorded with a photon energy of 55.5 eV.

RAS spectra show a typical line-shape characteristic for the (2×4) surface, as observed before.⁶ Photoemission spectra of the clean InP surface were taken for different photon energies. Curve fitting using Voigt functions is used to decompose the spectra into their components. In the case of the P $2p$ peak, we found a surface core-level emission shifted by 0.36 eV towards lower binding-energies with respect to the bulk emission line, as shown in Fig. 1. However, the Gaussian width of the spectrum is much larger than that of the P $2p$ spectrum of the phosphorous cap after desorption of the arsenic layer. This shows the presence of a local work-function variation at the surface. In the case of In $4d$ (Fig. 2), three components are required to reproduce the spectra. Thus, we obtain two components besides the bulk emission line, one shifted to higher (+0.38 eV) and one shifted to lower (−0.43 eV) binding energies. The appearance of the In $4d$ surface component at higher binding energy (+0.38 eV) and of the P $2p$ component at lower energy (−0.36 eV) can most likely be explained by a charge transfer from indium to phosphorous, as expected for a III–V-surface structure satisfying the electron counting rule.^{7,8} Accordingly, at the surface In atoms with one empty and P atoms with one doubly occupied dangling bond orbital should be present. The In $4d$ component shifted to lower binding energies could be interpreted as a second In $4d$ surface component, indicating the presence of other In bonding sites at the surface. Alternatively, this component could stem from small In metal clusters on the InP surface.

After deposition of CdS on the surface, only the bulk component of the P $2p$ core level remains (Fig. 1). The lack of any other P $2p$ emission lines in the spectra shows that no chemical compound formation takes place between the surface P and the deposited CdS. In contrast, the In $4d$ levels (Fig. 2) clearly evidence such a chemical reaction. Upon CdS deposition, the surface induced emission disappears and a new peak appears shifted by (0.45 ± 0.05) eV towards higher binding energy. The chemical interface reaction continues during deposition of CdS up to a nominal coverage of approximately 6 nm, as indicated by the increasing intensity of the “reacted” component with respect to the bulk emission line. The same kind of interface reaction was found by Wilke *et al.*¹ for the CdS/InP(110) interface. In analogy, the additional component can be attributed to an indium–sulfur reaction. Thus, the CdS/InP(001) interface is not an example of an abrupt heterostructure. Since an interfacial In_xS_y compound of several nm thickness is formed. S $2p$ or Cd $4d$ core levels have not been analyzed. The S $2p$ core level is out of the accessible energy range of the beam line. The Cd $4d$ peak, on the other hand, has a valence-band character, preventing a meaningful curve fitting analysis.

The RAS spectra plotted in Fig. 3 show a completely different line shape after deposition of a small amount of CdS on the InP surface. New features appear at 3.2 and 4.6 eV close to the energies of the E_1 , $E_1 + \Delta_1$, and E'_0 gaps of InP. That the bulk features are visible for even high CdS coverage, that the line shape remains essentially the same,

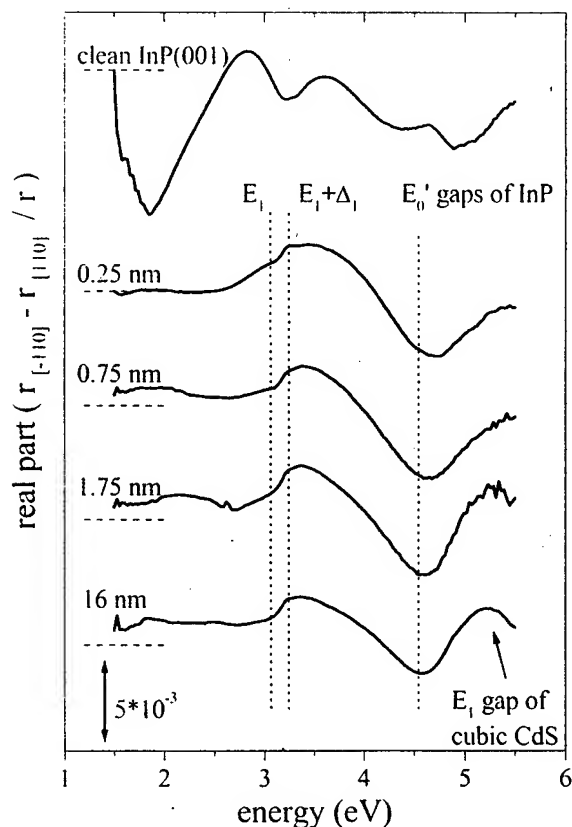


FIG. 3. Real parts of RAS spectra for the clean InP(001) surface (at the top) and increasing CdS coverage.

and that the feature around E'_0 is very broad, indicate that the CdS layer is not continuous, which is likely due to the interface reaction and a rough interface. From simulations, we can conclude that the line shape of the spectra are not simple but must have at least two contributions. After a few nm layer thickness, another weak feature appears around 5.3 eV, which is close to the energy of the E_1 gap of cubic CdS, indicating a cubic growth of CdS on the InP(001) surface (hexagonal CdS has two features in this spectral range). To verify the zinc-blende modification of CdS, Raman spectra (Fig. 4) were taken of CdS grown on InP(001). The dominant peaks in the spectra are clearly at the longitudinal-optical (LO) and the 2LO frequency positions at 305 and 610 cm^{-1} of CdS, respectively, as typical for cubic CdS.² These modes cannot be used to distinguish between the cubic and the wurtzite phase of CdS. However, for the naturally preferred wurtzite modification of CdS, additionally several strong modes in the range between 230 and 260 cm^{-1} are to be expected.⁹

B. Band lineup at the CdS/InP(001) interface

The valence-band spectra of the clean InP(001) surface as well as for different overlayers of CdS taken at a photon energy of 55 eV are shown in Fig. 5. Observable is a shift of the valence-band maximum with increasing CdS coverage. After deposition of a small amount of CdS, the shift of the valence-band maximum is mainly due to band bending

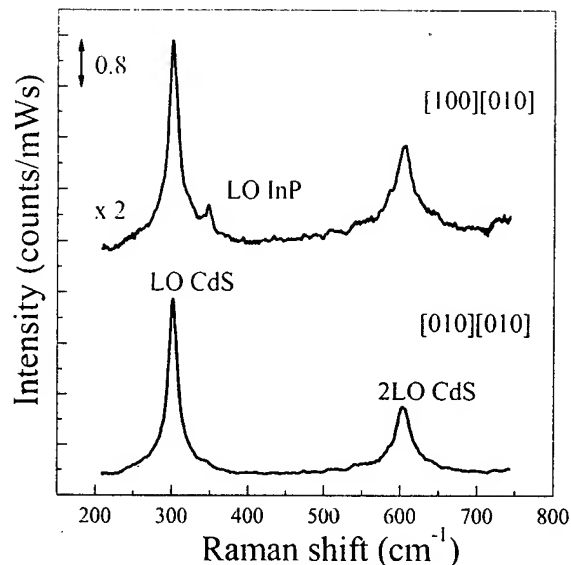


FIG. 4. Raman spectra of CdS/InP(001) in crossed and parallel configuration using the 488.0 nm (2.54 eV) emission line of an Ar^+ ion laser.

changes at the surface. Figure 6(a) shows the binding energy shifts of the In 4d bulk emission. The band bending saturates after about a 6 nm layer thickness, i.e., after completion of the $\text{In}_x\text{-S}_y$ interface reaction. In order to determine the valence-band offset ΔE_V , it is necessary to take the band

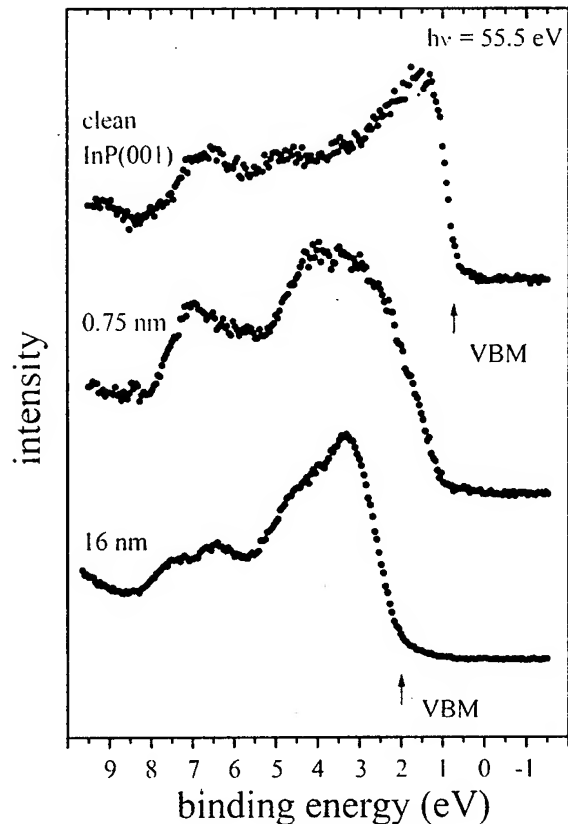


FIG. 5. Valence band of the clean InP(001) surface and for different thicknesses of CdS, recorded with a photon energy of 55.5 eV.

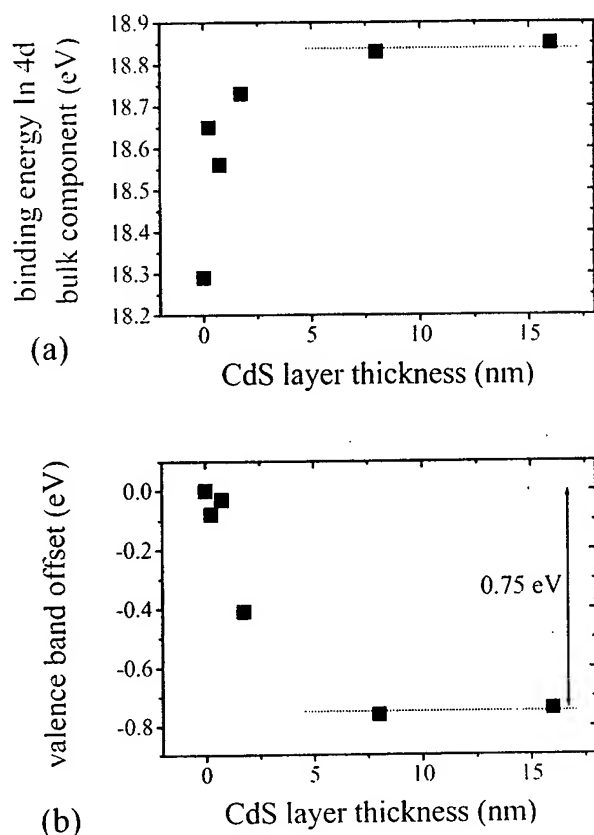


FIG. 6. Development of the binding energy of the In 4d bulk component (a) and the valence-band offset (b) with increasing CdS coverage.

bending into account by calibrating the valence-band maximum with respect to the In 4d bulk emission. The development of the valence-band offset in dependence on the CdS layer thickness is shown in Fig. 6(b). The final value for ΔE_V is reached above a nominal layer thickness of approximately 6 nm. This can be explained by the chemical reaction extending up to approximately 6 nm, which consumes most of the deposited S for the In_xS_y compound formation. Furthermore, in agreement with the RAS data, due to the strong chemical reaction, most likely the CdS layer does not grow in a perfectly laminar mode, at least for small coverage in the nm range. After a nominal CdS coverage of more than 6 nm, the final value for the valence-band offset of (-0.75 ± 0.1) eV is reached. The band alignment is of straddling type (Fig. 7).

The fundamental band gaps of InP [1.34 eV (Ref. 10)] and cubic CdS [2.36 eV (Ref. 11)] lead directly to a conduction-band discontinuity of $\Delta E_C = 0.27$ eV. To compare our results with previous experimental data, it is important to notice that all known data are based on a CdS/InP heterojunction with other substrate orientation. Our results are in very good agreement with photoemission and electron transport data of the CdS/InP(110) interface.^{1,12} Therefore, for the CdS/InP heterojunction, the band offsets seem not to depend on the crystal orientation of the interface, which may be a consequence of the strong interface reaction. More previous $C-V$ measurements by Shay *et al.*¹³ for CdS grown on

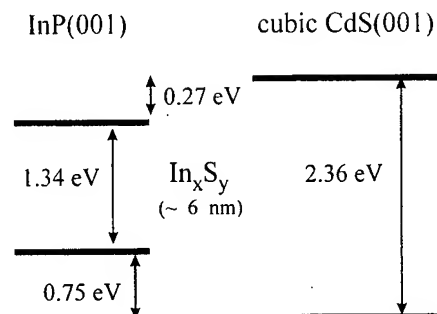


FIG. 7. Schematic band structure of the CdS/InP(001) heterojunction.

the B (phosphorous) face of InP(111) cleaned by chemical methods, yielded to a conduction band offset of 0.56 eV, according to a valence-band offset of 1.63 eV. However, in these experiments, the surface of the InP substrate was not controlled. Therefore, the disagreement of the latter in comparison to the photoemission experiments for (100) and (110) interfaces can most likely be a consequence of the (111) surface preparation.

IV. CONCLUSION

The growth of CdS on InP(001) was analyzed *in situ* by soft x-ray photoemission spectroscopy and reflectance anisotropy spectroscopy. A strong chemical reaction between indium and sulphur is observed at the CdS/InP(110) interface by analyzing the core-level spectra. *In situ* RAS spectra and separate Raman measurements reveal CdS in the metastable cubic modification. The valence-band offset is obtained with $\Delta E_V = (0.75 \pm 0.10)$ eV (straddling type). Similar results were previously obtained for the CdS/InP(110) interface, indicating that orientation specific effects seem to be negligible in the case of the CdS/InP heterostructure, probably due to the disruption of the interface by In_xS_y compound formation.

ACKNOWLEDGMENT

The authors would like to thank the Bundesministerium für Bildung und Forschung (05 622KTA 0) for financial support.

¹W. G. Wilke, R. Seedorf, and K. Horn, J. Vac. Sci. Technol. B 7, 807 (1989).

²D. R. T. Zahn, Ch. Maierhofer, A. Winter, M. Reckzügel, R. Srama, A. Thomas, K. Horn, and W. Richter, J. Vac. Sci. Technol. B 9, 2206 (1991).

³D. R. T. Zahn, Ch. Maierhofer, A. Winter, M. Reckzügel, R. Srama, U. Rossow, A. Thomas, K. Horn, and W. Richter, Appl. Surf. Sci. 56-58, 684 (1992).

⁴A. D. Katnani and G. Margaritondo, Phys. Rev. B 28, 1944 (1983).

⁵*Epiptic, Linear and Nonlinear Optical Spectroscopy of Surfaces and Interfaces*, edited by J. McGilp, D. Weaire, and C. H. Patterson (Springer, Berlin, 1995).

⁶K. Knorr, M. Pristovsek, U. Resch-Esser, N. Esser, M. Zorn, and W. Richter, J. Cryst. Growth (to be published).

⁷M. M. Sung, C. Kim, H. Bu, D. S. Karpuzov, and J. W. Rabalais, Surf. Sci. 322, 116 (1995).

⁸C. D. MacPherson, R. A. Wolkow, C. E. J. Mitchell, and A. B. McLean, *Phys. Rev. Lett.* **77**, 691 (1996).

⁹R. M. Martin and T. C. Damen, *Phys. Rev. Lett.* **26**, 86 (1971).

¹⁰*Landolt-Börnstein New Series*, edited by O. Madelung, M. Schulz, and H. Weiss (Springer, Berlin, 1982), Vol. 17.

¹¹D. R. T. Zahn, G. Kudlek, U. Rossow, A. Hoffmann, I. Broser, and W. Richter, *Adv. Mater. Opt. Electron.* (to be published).

¹²Ch. Schultz, M. von der Emde, and D. R. T. Zahn, *Phys. Status Solidi A* **145**, 497 (1994).

¹³J. L. Shay, S. Wagner, and J. C. Phillips, *Appl. Phys. Lett.* **28**, 31 (1976).

Effect of growth conditions on surface roughening of relaxed InGaAs on GaAs

T. Pinnington, C. Lavoie,^{a)} and T. Tiedje^{b)}

Advanced Materials and Process Engineering Laboratory, Department of Physics and Astronomy,
University of British Columbia, Vancouver, BC, Canada V6T 1Z4

(Received 12 January 1997; accepted 22 April 1997)

We report elastic light scattering measurements of the surface morphology of strained $\text{In}_x\text{Ga}_{1-x}\text{As}$ on GaAs, grown by molecular beam epitaxy at different growth temperatures and In contents. During strain relaxation through formation of interfacial misfit dislocations, the surface of the film roughens in response to inhomogeneous surface strains produced by the interfacial misfit dislocations. The time dependence of this roughening is modeled by an Edwards–Wilkinson equation in which the deposition flux noise is neglected and the inhomogeneous surface stress is the only driving term. © 1997 American Vacuum Society. [S0734-211X(97)09204-4]

I. INTRODUCTION

Interface roughness in semiconductor heterostructures is an important factor in the performance of various kinds of electronic and optoelectronic devices. It can limit the mobility of electrons in field effect transistors, scatter light in optical waveguides, and broaden the gain spectrum of quantum well lasers.¹ Control of the structure of semiconductor surfaces on a submicron length scale during growth is needed in the fabrication of gratings for distributed feedback lasers and quantum wire and quantum dot structures. Surface and interface morphology in semiconductor heterostructures is derived from the surface structure of epitaxial thin films during growth and is affected by a variety of phenomena, including strain relaxation through formation of misfit dislocations and kinetic roughening due to the random flux of depositing atoms.²

Elastic light scattering is one of the simplest methods for measuring surface morphology during film growth. It is relatively noninvasive and is highly sensitive to small changes in surface structure on in-plane length scales, on the order of an optical wavelength. Typical polished surfaces of substrates used in epitaxial growth are smooth on nearly the atomic scale. In this case, the background scattering from the starting surface is very small and it is possible to detect changes in the surface structure with a submonolayer amplitude by light scattering. The intensity of the diffusely scattered light from a rough surface is proportional to the power spectral density of the surface morphology, provided that the surface is not too rough. In this case, the light scattering measures the power spectral density of the surface roughness at a spatial frequency that is defined by the wavelength of the light, the angle of incidence, and the scattering angle, as explained elsewhere.³

In this article, we describe the time evolution of the surface roughness of strained $\text{In}_x\text{Ga}_{1-x}\text{As}$ films grown on (100)

GaAs by molecular beam epitaxy. The surface roughness is determined from *in situ* elastic light scattering measurements during growth. We compare the time dependence of the surface roughness with the predictions of an Edwards–Wilkinson type model in which the surface roughening is driven by the inhomogeneous surface strain generated by interfacial misfit dislocations. This model can reproduce the spatial frequency dependence and time evolution of the surface morphology for this materials system, as reported elsewhere.⁴ In this article, we use the model to interpret new measurements of the crystallographic anisotropy of the surface roughening, and to examine for the first time the effect of growth temperature and indium content on the morphology dynamics.

II. EXPERIMENT

The light scattering measurements were carried out in a VG V80H molecular beam epitaxy system as reported earlier.^{5,6} Two different angles of incidence were used for the 488 nm light from an air-cooled Ar ion laser. In one case, the light was incident normal to the surface of the substrate through the pyrometer port, with silicon pin diode detectors at an effusion cell port or at a shutter port, in which case the scattered light intensity is sensitive to the spatial frequencies of 5.4 and 10.5 μm^{-1} in the plane of the substrate. Alternatively, the light was incident through an effusion cell port with the scattered light detected at the pyrometer port or in a backscattering geometry from the shutter port next to the effusion cell port used for the incident light. In this case, the relevant spatial frequencies in the surface of the substrate are 5.4 and 16 μm^{-1} , respectively. The spot size of the laser on the surface of the wafer was about 2 mm in diameter. The location of the beam spot on the GaAs (100) nominally on-axis substrates ($\pm 0.5^\circ$), was selected to avoid visible surface defects but was not subsequently moved during film growth. In all the ports, optical access was obtained via mirrors in the vacuum chamber so that there is no direct line of sight from the hot substrate to the window, in order to minimize window coating. The diffusely scattered light was col-

^{a)}Present address: IBM T.J. Watson Research Center, Yorktown Heights, NY.

^{b)}Also with Dept. of Electrical and Computer Engineering; Electronic mail: tiedje@physics.ubc.ca

lected with a 60-mm-diam lens, passed through a laser line filter, and detected with a silicon pin diode and lock-in detection system.

From the repeat growths of GaAs buffer layers under nominally the same growth conditions, we believe that the reproducibility of the intensity of the scattered light between runs is not better than a factor of about two, due to variations in optical throughput. The time dependence of the scattered light signal during growth, on the other hand, is found to be quite consistent between runs. Over a period of many growths, particularly for growths that involve high substrate temperatures at which significant Ga or In desorption takes place (e.g., 700 °C for GaAs or 600 °C for InGaAs), window and mirror coating is a problem, which is solved by periodically removing and cleaning these optical elements.

With light incident through the normal incidence pyrometer port, it is possible to simultaneously collect light scattered in two orthogonal planes containing the $[110]$ and $[1\bar{1}0]$ directions in the substrate surface during growth. Above a critical thickness, strain relaxation takes place through formation of misfit dislocations. The surface of the InGaAs then develops the well-known cross-hatch surface morphology with two orthogonal one-dimensional arrays of randomly spaced ridges. The angular distribution of the scattered light can be thought of as the spatial Fourier transform of the morphology of the scattering surface. Therefore the light is scattered primarily in two orthogonal planes at right angles to the cross-hatch surface roughness. In order to detect this scattering, the sample has to be very carefully aligned before growth so that the two orthogonal detectors lie in planes containing the $[110]$ and $[1\bar{1}0]$ directions of the sample. This is done using the cleaved edges of the quarter wafer substrates as direction references.

The substrate temperature was measured by optical band gap thermometry with an estimated absolute accuracy of about ± 10 °C, and relative accuracy of about ± 2 °C.⁷ The composition of the films was estimated initially from the relative fluxes of Ga and In, measured with a retractable ion gauge placed in the substrate position. X-ray diffraction and optical transmission measurements were then carried out on a number of samples to calibrate the In contents from the flux measurements. The In contents are believed to have an uncertainty of about 10%. The As_2 flux as measured with an ion gauge, was maintained constant at a As_2 :Ga ratio of 3.5:1, in all the runs. GaAs substrates were prepared in an ultrahigh vacuum (UHV) by thermal evaporation of the surface oxide followed by the growth of a 0.5 μm (typical) GaAs buffer layer to smooth the surface.

In Fig. 1, we show the scattered light intensity as a function of time during growth for light at normal incidence, for the two scattering planes containing the $[110]$ and $[1\bar{1}0]$ directions. The figure shows that the surface roughening along $[110]$ is significantly larger than along $[1\bar{1}0]$. Both directions show a threshold for the onset of surface roughening. This threshold is a critical thickness for the creation of misfit dislocations at the InGaAs/GaAs interface. The threshold thickness for the $[110]$ direction is slightly larger than for the

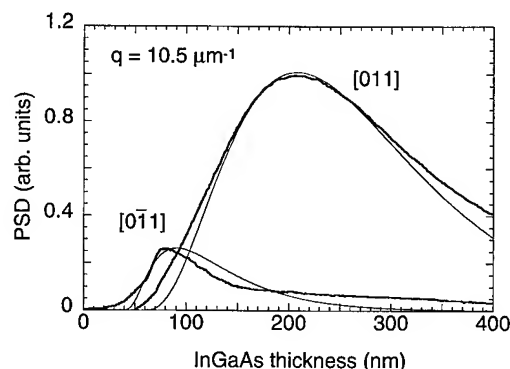


Fig. 1. Time dependence of the scattered light intensity at a spatial frequency of $10.5 \mu\text{m}^{-1}$ in planes parallel to the $[110]$ and $[1\bar{1}0]$ directions for an $\text{In}_{0.18}\text{Ga}_{0.82}\text{As}$ film grown at 490 °C, at $1 \mu\text{m/h}$. The two experimental curves were measured during the same growth and show the observed relative values of the scattered light intensities. The thin lines are the calculated surface power spectral densities as a function of time during growth with $\nu = 0.17 \times 10^{-12}$ and $1.4 \times 10^{-12} \text{ cm}^2/\text{s}$ and relaxation thickness of 60 and 40 nm for scattering in the $[110]$ and $[1\bar{1}0]$ directions, respectively. The magnitude of the constant A for the scattering along $[110]$, has been reduced by a factor of 5.8 relative to the other scattering direction in order to match the data.

$[1\bar{1}0]$ direction. This is consistent with atomic force microscopy images which show that the dislocation lines start parallel to $[110]$ first.⁵ The weak oscillation in the scattered intensity at the beginning of the growth of the InGaAs, is believed to be a thin film interference oscillation due to the different index of refraction of the InGaAs as compared with the GaAs substrate.

In Figs. 2 and 3, we show the time evolution of the scattered light intensity for two different growth temperatures, 452 and 515 °C, at spatial frequencies of 16 and $5.4 \mu\text{m}^{-1}$, respectively. The critical thickness appears to decrease with increasing temperature. This result is to be expected if the misfit dislocation formation is kinetically limited; as the temperature is increased, the elastic strain and misfit density will

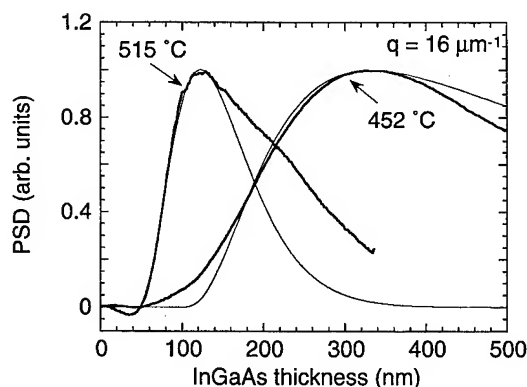


Fig. 2. Time dependence of the scattered light intensity along $[110]$ during growth of InGaAs on GaAs at 515 and 452 °C for a spatial frequency of $16 \mu\text{m}^{-1}$ (heavy lines). The thin lines are the calculated surface power spectral densities at $16 \mu\text{m}^{-1}$ for $\nu = 0.2 \times 10^{-12}$ and $0.008 \times 10^{-12} \text{ cm}^2/\text{s}$ for 515 and 452 °C, respectively. The relaxation thicknesses are 45 (515) and 100 nm (452 °C) for the model. Both the model and the data have been normalized to unity at the peaks.

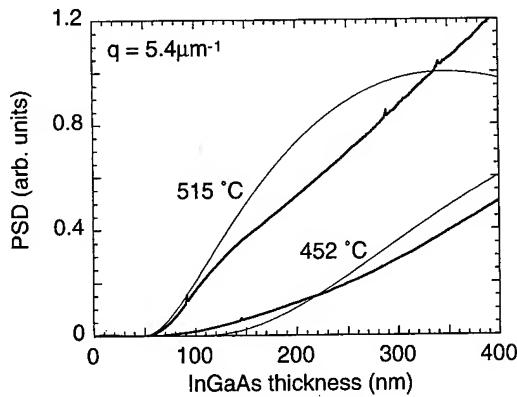


FIG. 3. Time dependent light scattering data and model curves as in Fig. 2, except that the scattering geometry corresponds to a spatial frequency of $5.4 \mu\text{m}^{-1}$. The model curves have been normalized to unity at their peak values, which is offscale in the case of the model for the 452 °C sample. The amplitude of the experimental data (heavy lines) has been scaled to match the model.

be closer to the equilibrium values. Also we note that the high spatial frequency light scattering has a peak as a function of time whereas the low spatial frequency data continues to increase over the entire range. In Fig. 4, we show a similar experiment for three different In concentrations. Notice that the thickness for the onset of roughening (i.e., the critical thickness), increases as the In content and lattice mismatch is reduced, as expected.

All of the experimental data have been normalized to the background scattering intensity from the GaAs buffer layer. Then all of the experimental curves in Figs. 1–4 are offset by one so that they start at zero on the vertical axis. The justification for this procedure is as follows. In contrast to the intense one-dimensional scattering lines associated with the strain relaxation, the diffuse background scattering from the buffer layer is approximately isotropic and remains nearly constant throughout the subsequent InGaAs growth, as verified by *ex situ* light scattering analysis of films grown to

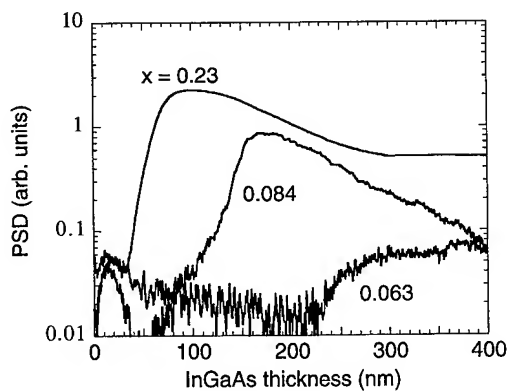


FIG. 4. Time dependence of the scattered light intensity along $[1\bar{1}0]$ at $q = 16 \mu\text{m}^{-1}$ during InGaAs growth with three different In concentrations as indicated. The growth temperature is 490 °C. The experimental data has been scaled to the background scattering from the buffer layer prior to the start of the growth of the InGaAs, but is not otherwise scaled.

different thicknesses.⁶ During the *in situ* measurement, some of this diffuse background intensity is collected by the detector lens along with the sharp scattering line from the surface cross hatch. The one-dimensional scattering of interest can be isolated by subtracting the diffuse background signal, as approximated by the signal intensity immediately following the buffer layer growth.

III. MODEL

In order to explain the observed time dependence of the scattered light intensity during growth, we use a model developed earlier.⁴ We provide a brief outline of the model here for clarity. In this model, the surface roughening during the InGaAs growth is produced by surface migration in response to surface strain gradients, which result from the inhomogeneous strain produced by the misfit dislocations at the film-substrate interface. The surface roughening of interest has the form of one dimensional arrays of ridges parallel to the two orthogonal systems of misfit dislocations. Thus the roughening of interest is one dimensional, which makes the mathematical modelling of the roughening relatively simple.

Atoms migrate on the surface of the InGaAs in response to elastic-stress-related chemical potential gradients. The contribution of the surface strain to the chemical potential $\mu(x, t)$ produced by the local surface strain field $\epsilon(x, t)$ at position x and time t is equal to

$$\mu(x, t) = \frac{Y\Omega}{2} \epsilon^2(x, t), \quad (1)$$

where Y is the biaxial modulus and Ω is the volume of a unit cell. This expression can be linearized as follows

$$\mu(x, t) \approx \frac{Y\Omega}{2} \left[-\bar{\epsilon}^2(t) + 2\bar{\epsilon}(t)f + 2\bar{\epsilon}(t)\delta_\epsilon(x, t) \right], \quad (2)$$

in the case where the local deviation in the strain field, $\delta_\epsilon(x, t)$, at the surface due to the misfit dislocations is small compared to the average strain $\bar{\epsilon}(t)$ in the lattice-mismatched overlayer; f is the initial in-plane strain due to the lattice mismatch. The contribution of shear strain produced by the mixed screw/edge dislocations is of order $(\delta_\epsilon)^2$ and can be neglected. The flux of material on the surface of the film in response to the chemical potential gradient is equal to

$$j = -\frac{Dn}{kT} \nabla \mu(x, t). \quad (3)$$

In this equation, D is the surface diffusion coefficient and n is the concentration of diffusing atoms on the surface.

In order to determine the time dependence of the surface morphology during growth, we need an equation which describes the dynamical behavior of the surface morphology. The simplest equation which has the required symmetry is the Edwards–Wilkinson equation.² In this case, there is a surface-gradient-driven atom current equal to $\nu \nabla h(x, t)$ where $h(x, t)$ is the surface height and ν must be greater than zero for stable growth. The resulting equation for the surface morphology is

$$\frac{\partial h(x,t)}{\partial t} - \nu \nabla^2 h(x,t) = \frac{\Omega D n Y}{kT} \bar{\epsilon}(t) \nabla^2 [\delta_\epsilon(x,t)]. \quad (4)$$

The microscopic physical mechanism behind the ∇^2 term on the left hand side of Eq. (4) is not known for InGaAs surfaces, however two candidate mechanisms are evaporation of atoms (indium?) from the surface and diffusion bias produced by step-edge potential barriers (Schwoebl barriers). The contribution of the surface curvature to the chemical potential has been neglected. To include the effect of the curvature one would need to add a $\nabla^4 h$ term² to Eq. (4).

We assume that the misfit dislocations in the one-dimensional arrays are randomly spaced with average linear density $\rho(t)$. In this case, the Fourier transform of $\delta_\epsilon(x,t)$ is

$$\hat{\delta}_\epsilon(q,t) = \sqrt{2\rho(t)} \hat{\epsilon}_1(q,t), \quad (5)$$

where $\hat{\epsilon}_1(q,t)$ in Eq. (5) is the Fourier transform of the surface strain distribution due to a single isolated misfit dislocation at the film-substrate interface and q is spatial frequency. The film thickness is $z = gt$, where g is the growth rate. An analytical expression is available for the strain field $\epsilon_1(x,z)$ at the surface due to misfit dislocations for the system of interest, namely InGaAs, deposited on (100) GaAs.⁸ The dislocations are mixed screw-edge dislocations with a Burgers vector at 60° with respect to the line direction. For a dislocation of this type located at $x=0$ and $z=0$, with a Burgers vector of magnitude b , the strain at the surface is given by

$$\epsilon_1(x) = \frac{-bz}{\pi} \frac{(x^2 - \sqrt{2}xz)}{(x^2 + z^2)^2} \Leftrightarrow \frac{b}{2} e^{-z|q|} (z|q| - 1 + i\sqrt{2}z|q|), \quad (6)$$

where \Leftrightarrow denotes Fourier transform. This expression does not include the shear strain.

Solving the Fourier transform of Eq. (4), we find:

$$\hat{h}(q,z) = A q^2 e^{-\nu q^2 z/g} \times \int_0^{z/z_r-1} e^{\nu q^2 \zeta z_r/g} \bar{\epsilon} \rho^{1/2} \hat{\epsilon}_1[q, z_r(\zeta+1)] d\zeta, \quad (7)$$

$$A = \frac{\sqrt{2} z_r \Omega^2 D n Y}{g k T}, \quad (8)$$

where g is the growth rate and z_r is the thickness at which the film begins to relax. We have used a $1/z$ dependence for the residual strain in the film for thicknesses greater than z_r . This is equivalent to the Matthews-Blakeslee model⁹ if a logarithmic term is neglected, and is consistent with recent *in situ* measurements of strain during growth of relaxed InGaAs on GaAs.¹⁰ The average dislocation density, ρ , at each thickness is taken to be linearly related to the relieved strain through the in-plane component of the Burgers vector: $\rho = [2(f - \bar{\epsilon})/b]$. For $z < z_r$, $\rho = 0$.

As pointed out above, the scattered light intensity is proportional to the power spectral density of the surface which in our case is equal to $h^2(q,t)$. This can be obtained numerically by integration of Eq. (7). In the numerical simulations

of the power spectral density, we use the experimental values for the growth rate ($1 \mu\text{m/h}$) and the spatial frequency, defined by the scattering geometry and the optical wavelength. The lattice mismatch is obtained from the In content of the films. This leaves three fitting parameters, namely the constant A , the parameter ν from the Edwards-Wilkinson equation, and the thickness z_r for the onset of misfit dislocations.

IV. DISCUSSION

In Fig. 1, we show fits to the scattering along the orthogonal $[110]$ and $[\bar{1}\bar{1}0]$ directions, measured simultaneously during growth of an 18% In content film, and showing the observed relative scattering intensities. The experimental data is fit using ν and z_r as fitting parameters, with the amplitudes scaled to match the data. Although the model predicts the scattering along $[110]$ to be larger than in the orthogonal direction, the ratio of the two calculated peak intensities is a factor of 5.8 larger than is observed in the experimental data. The fits also show that the critical thickness is larger for the dislocations perpendicular to $[110]$ and that the parameter ν , which describes how the surface smooths in response to a perturbation, is smaller in the $[110]$ direction than in the $[\bar{1}\bar{1}0]$ direction. As pointed out above, this difference in the onset of misfit dislocations has been observed in other experiments as well.⁵ The critical thicknesses in both directions are significantly larger than the equilibrium critical thickness as one would expect for measurements made during growth, due to the kinetic constraints on the formation of misfit dislocations. This discrepancy with the equilibrium value is well known in the literature. The smaller value of ν along $[110]$ is consistent with the crystallographic asymmetry in (100) zincblende surfaces. We note that the high ν direction ($[\bar{1}\bar{1}0]$) coincides with the high diffusion coefficient direction on GaAs (elongated mounds in this direction).

In Figs. 2 and 3, we show fits to the light scattering data at two different temperatures and spatial frequencies. In this case, we have scaled the peak values to unity. The high spatial frequency data ($16 \mu\text{m}^{-1}$) shows a peak for the film thickness range presented while in the case of the low spatial frequency data ($5.4 \mu\text{m}^{-1}$), the data increases monotonically over the entire range. The peak in the high spatial frequency data shifts to smaller thicknesses at higher temperatures. The model can reproduce these effects if the relaxation thickness and ν parameter are allowed to be temperature dependent. In order to fit the data, the ν parameter must increase from 0.008×10^{-12} to $0.2 \times 10^{-12} \text{ cm}^2/\text{s}$ and the relaxation thickness decreases from 100 to 45 nm as the temperature goes up from 452 to 515 °C. The decrease in the thickness at which dislocations start to form at high temperatures is reasonable given that the observed threshold thickness is large compared with the equilibrium value. We expect kinetic bottlenecks to be reduced at high temperatures.

Since the physical process that determines ν is not known, it is difficult to make predictions about its temperature dependence. Assuming that the temperature dependence arises from a thermally activated process, we find an activation energy of 2.5 eV from the two values for ν that we have

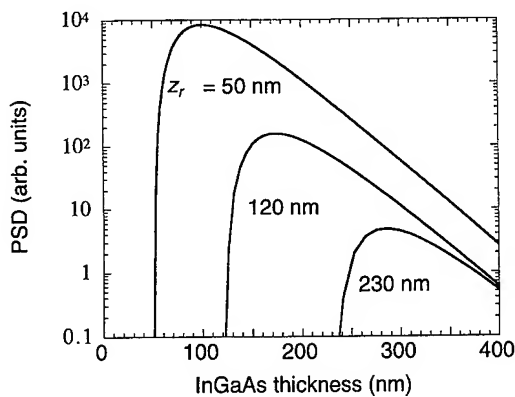


FIG. 5. Time dependence of the power spectral density calculated from the model discussed in the text for $q = 16 \mu\text{m}^{-1}$ with three different values of the relaxation thickness as indicated, designed to match the data in Fig. 4, with $\nu = 0.6 \times 10^{-12} \text{ cm}^2/\text{s}$.

obtained at 452 and 515 °C. This activation energy is large compared with what one might expect for surface diffusion processes.

In Figs. 4 and 5, we show calculated and experimental roughness data during growth for three different In contents, all measured at the same temperature. In this case, we have retained the relative amplitude information, as the relative amplitudes of the experimental curves differ by more than the experimental uncertainty. All of the experimental data in Fig. 4 were first normalized to the background scattering from the buffer layer at the start of InGaAs growth and then the background scattering was subtracted. In the case of the 6.3% In sample in Fig. 4, a background of slightly less than unity was subtracted.

The model curves in Fig. 5 were calculated with $\nu = 0.6 \times 10^{-12} \text{ cm}^2/\text{s}$, which was selected to match the data for the sample with a 8.4% In content. The thickness for the onset of misfit dislocations was used as a fitting parameter. The difference in the relaxation thicknesses for the 6.3% and 8.4% samples is large compared with what one would expect from the difference in the nominal In contents. Although the model can reproduce the general trend of lower scattering at lower In concentrations and larger critical thicknesses, the model predicts a much larger change in the intensity than is

observed experimentally. This could be explained by a change in the prefactor A , (which depends on the surface diffusion) with In concentration, or by a change in ν , or both.

V. CONCLUSION

In conclusion, we report light scattering measurements of surface roughening during relaxation of strained $\text{In}_x\text{Ga}_{1-x}\text{As}$ on (100) GaAs. Surface roughness measurements are made for different temperatures and In concentrations during growth of $\text{In}_x\text{Ga}_{1-x}\text{As}$ on GaAs (100). An Edwards–Wilkinson type model is able to account for the observed behavior of the surface morphology. In this model, strain gradients from misfit dislocations act as a driving term for surface diffusion. The parameter ν in the Edwards–Wilkinson equation is found to increase with temperature and depend on crystal orientation. This work shows that light scattering can provide detailed information on the dynamical behavior of surface morphology during thin film growth.

ACKNOWLEDGMENTS

The authors thank E. Nodwell and B. Haveman for help with the measurements and K. Kavanagh and S. Patitsas for useful discussions.

¹J. A. Dura, J. G. Pellegrino, and C. A. Richter, Appl. Phys. Lett. **69**, 1134 (1996); D. K. Fork, F. Armani-Leplingard, M. Lui, and R. A. McFarlane, J. Lightwave Technol. **14**, 611 (1996); W. G. Bi and C. W. Tu, J. Appl. Phys. **78**, 2889 (1995).

²A. L. Barabasi, *Fractal Concepts in Surface Growth* (Cambridge University Press, Cambridge, England, 1995).

³E. L. Church, H. A. Jenkinson, and J. M. Zavada, Opt. Eng. (Bellingham) **16**, 360 (1977).

⁴T. Pinnington, C. Lavoie, T. Tiedje, B. Haveman, and E. Nodwell, Phys. Rev. Lett. (to be published).

⁵C. Lavoie, T. Pinnington, E. Nodwell, T. Tiedje, R. S. Goldman, K. L. Kavanagh, and J. L. Hutter, Appl. Phys. Lett. **67**, 3744 (1995).

⁶C. Lavoie, Ph.D. thesis, University of British Columbia, 1994.

⁷S. R. Johnson, C. Lavoie, T. Tiedje, and J. A. Mackenzie, J. Vac. Sci. Technol. B **11**, 1007 (1993).

⁸F. Jonsdottir and L. B. Freund, Mech. Mater. **20**, 337 (1995).

⁹J. W. Matthews, S. Mader, and T. B. Light, J. Appl. Phys. **41**, 3800 (1970).

¹⁰P. Moeck, B. K. Tanner, C. R. Whitehouse, A. G. Cullis, G. Lacey, G. F. Clarke, B. Lunn, J. C. H. Hogg, A. M. Keir, A. D. Johnson, G. W. Smith, and T. Martin, *Proceedings of the 23 International Conference on the Physics of Semiconductors*, edited by M. Scheffler and R. Zimmerman (World Scientific, Singapore, 1996), Vol. 2, p. 927.

In-rich 4×2 reconstruction in novel planar growth of InAs on GaAs(001)

Q. K. Xue,^{a)} Y. Hasegawa, T. Ogino, H. Kiyama, and T. Sakurai
Institute for Materials Research, Tohoku University, Sendai 980-77, Japan

(Received 22 January 1997; accepted 16 April 1997)

Molecular beam epitaxy growth of lattice mismatched InAs/GaAs(001) system is studied by *in situ* scanning tunneling microscopy (STM) and reflection high energy electron diffraction. We found that deposition of submonolayer (~ 0.6 ML) In on the GaAs(001)-As-rich 2×4 surface could result in a new well-ordered 4×2 reconstruction, and that if the growing front maintains this reconstruction, the multilayer InAs will grow two-dimensionally and usually observed 3D islanding is completely suppressed. Atomic structures for the 4×2 reconstruction are discussed on the basis of voltage-dependent STM image. A "domain wall" structure, representing a new type of surface strain relief mechanism in the novel growth will also be discussed. © 1997 American Vacuum Society. [S0734-211X(97)09004-5]

I. INTRODUCTION

"Self-organization" in heteroepitaxially lattice-mismatched systems, arising from strain-induced coherent 3D islanding has been of great importance for fabrication of 0D confined nanostructures.¹⁻⁶ In traditional "quantum wells" structure where electrons are confined only in one direction, one expects that the Stranski-Krastanov (2D-3D) morphology transition can be delayed or suppressed, a more planar growth and thus both morphologically and compositionally abrupt interfaces can be achieved, which are crucial for optimum performance of electronic and optoelectronic devices based on superlattice and heterojunction.⁷⁻¹⁴ For the InAs/GaAs with a lattice mismatch of 7.2%, in order to reduce the accumulated epitaxial strain energy in the grown layer, the 2D-3D transition occurs at approximately 2 MLs (the so-called critical thickness) under standard molecular beam epitaxy (MBE) growth conditions.

MBE growth of any epitaxial film is inherently a nonequilibrium process, the morphology of the growing film and its evolution are governed by an interplay between kinetic and thermodynamic aspects of growth.¹¹⁻¹⁴ This provides us an opportunity to control the growth mode by manipulating the growth conditions and kinetics. One of the examples is that by using extremely In-rich growth conditions, the usually observed 3D morphology roughening could be significantly suppressed.⁷⁻¹⁰ Schaffer *et al.* have reported a basically 2D growth even for an epilayer thickness of 1800 Å under such conditions,⁷ which has resulted in superior interface quality and remarkably improved photoluminescence intensity.⁸ For the unusual growth, several questions are raised: (1) what does the surface morphology really look like? (2) what are surface structure/reconstructions? (3) what type of strain relaxation mechanism works and how does it work? (4) is it due to limited kinetics or energetically favorable? In the present study, by using a high-performance *in situ* molecular beam epitaxy-scanning tunneling microscope (MBE-STM) system, we have investigated the evolution of the surface

morphology and structure as a function of the III-V flux ratio and epilayer thickness in the 2D growth of InAs/GaAs(001), aiming at understanding the growth at the atomic scale.

II. EXPERIMENT

The MBE-STM system used in the study and the preparation of the GaAs(001)- 2×4 substrates have been described elsewhere.¹⁵ The surface consists of characteristic As-dimer missing-dimer rows that are evenly separated by 16 Å along the direction perpendicular to the dimer row. By dimerization and dimer-missing, all the surface dangling bonds of As are filled with two electrons and those of Ga empty, assuring a stable semiconducting surface of the 2×4 . Two well-established atomic models are schematically shown in Fig. 1.^{16,17} The most ordered β phase dictated in the upper panel of Fig. 1 was used as the substrate for InAs growth. The starting InAs wetting layer was prepared by depositing submonolayer (approximately 0.6 ML) In on the As-rich 2×4 surface at a substrate temperature of 450 °C. After deposition, the samples were allowed to maintain at 450 °C for about 5 min. The surfaces prepared in this way display a very sharp 4×2 reflection high-energy electron diffraction (RHEED) pattern.⁷ The subsequent multilayer InAs growth was carried out by using migration enhanced epitaxy (MEE) to eliminate the kinetic complications. Our "quench-and-look" STM experiment reveals that the As/In flux ratio and shutter switching time in MEE are critical for the layer-by-layer mode and that a smooth surface cannot grow unless the resulting growth front reproduces the 4×2 symmetry in the wetting layer.¹⁸ These observations demonstrate the very special function of this surface serving as template for the planar growth.

III. RESULTS AND DISCUSSION

Shown in Fig. 2(a) are typical filled-states STM images of the 4×2 surface. Figure 2(a) was recorded at a sample bias of -2.2 V while Fig. 2(b) at -1.6 V. The cross-section profiles L1, L2, and L3 are shown in the lower panel of the figure. The STM image exhibits a high degree of perfection of this surface, without any kink defects. The surface con-

^{a)}Present address: Physics Department, North Carolina State University, Raleigh, North Carolina, 27695-8202; Electronic mail: xue@unity.ncsu.edu

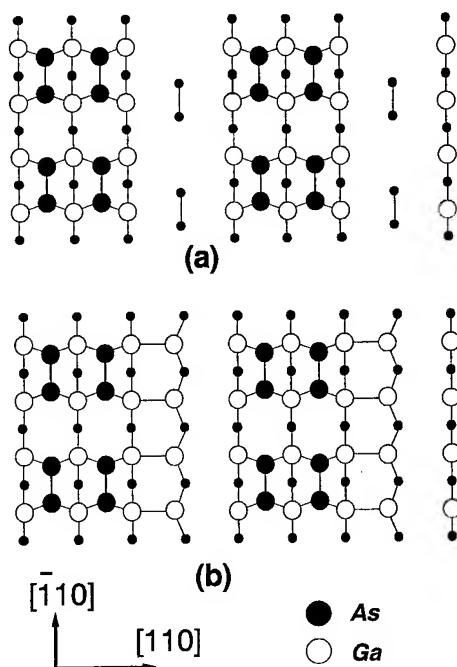


FIG. 1. Atomic models of the GaAs(001) As-rich 2×4 reconstruction. (a) Two dimer model for Ga-rich 2×4 by Farrell *et al.* (Ref. 16), (b) two dimer model for As-rich 2×4 by Chadi (Ref. 17). The black solid circles correspond to the As atoms, while the open circles the Ga atoms.

sists of straight line and evenly spaced humps along the [110] direction with a unit cell consisting of $16.0 \text{ Å} \times 8.0 (\pm 0.23) \text{ Å}$ [the 4×2 unit cell is highlighted by a rectangle in Fig. 2(b)]. The 4× along the [-110] direction and the 2× along the [110] direction translational symmetries are more clearly indicated by the cross-section profiles L1 and L2, respectively. In these filled states images, we do not see much contrast change for the line and the hump. When the bias voltage becomes smaller, we observe a distinct 1× periodicity along the line built up by smaller humps, as indicated by the cross section profile L3. If studying the Fig. 2(b) more carefully, it is found that the small humps forming the 1× periodicity always straddle the larger 2× humps along the [110] direction, which are highlighted by the empty and solid circles, respectively, suggesting that the 2× and 1× humps should tunnel from different species. All these basic features are quite different from what observed for the homoepitaxially grown InAs(001) In-rich 4×2 surface,^{19–21} and also for the 2×4 substrate,¹⁵ therefore, they should have different structures. Because the 2× direction agrees with the In dimerization direction, we conclude that these 2× (larger) humps are due to tunneling from In dimers.

We also obtained the empty states STM images (not shown here). All the basic hump-plus-line features are exactly the same as those in the filled states image, but the 2× humps become brighter than the lines. This contrast change corresponds to our identification that the 2× humps are due to tunneling from the In atoms.²² Based on these observations, we assign the individual 2× humps to be tunneling from the first layer In dimer, the lines from the second layer As, and propose a tentative model for this new reconstruction in Fig. 3(a). According to this model, the twofold periodicity is due to the dimerization of the newly deposited In adatoms in the [110] direction, and the regular In-dimer missing gives rise to the fourfold periodicity along the [-110] direction. The exposed As adatoms in the missing rows dimerize to their neighbors and their uniform arrangement explains the 1× periodicity along the line as shown in Fig. 2(b). The model abides with the electron counting model and assures that all arsenic dangling bonds fully filled with two electrons and all In dangling bonds empty leave no net charge at the surface, corresponding to a stable semiconducting surface.²³ The model gives a surface In coverage of 0.5 ML, in good agreement with the experiment. The observed contrast differences between the filled- and empty-states images are accounted for by the different energy levels of the As dangling bonds and In dimer antibonding orbitals.^{22,23}

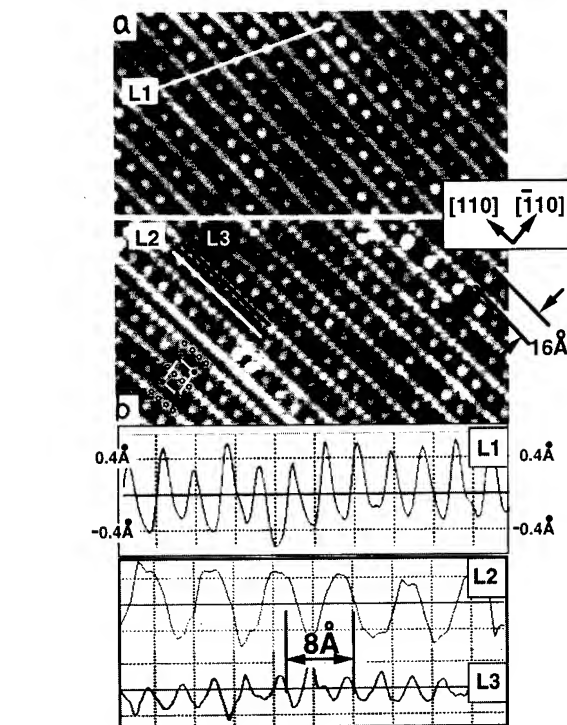


FIG. 2. Filled states STM images of the new In-terminated 4×2 surface. (a) $V_s = -2.2 \text{ V}$; (b) $V_s = -1.6 \text{ V}$. The small solid and open circles in (b) highlight the 2× and 1× humps, respectively, and their relative arrangement can be clearly seen in light of the superimposed unit cell. No thermal drift correction was made with the STM images, so the 4×2 unit cell looks like a parallelogram. The cross-section profiles L1, L2, and L3 are shown in the lower two panels of the figure.

Alternative models have been proposed previously for (001) III(Ga,In)-rich 4×2 surface which is also shown in Fig. 3. Figure 3(b) is the one-dimer model proposed by Ohkouchi and Ikoma²⁰ and supported by a very recent study.²¹ At a glance, an agreement between the one-dimer model and the STM image seems more reasonable than that for our model. A serious problem with this model is that one should expect a contrast difference of 3 Å (the bilayer step height of InAs) for the first and third layer In dimers, while the observed contrast is nearly absent in the image. There is no

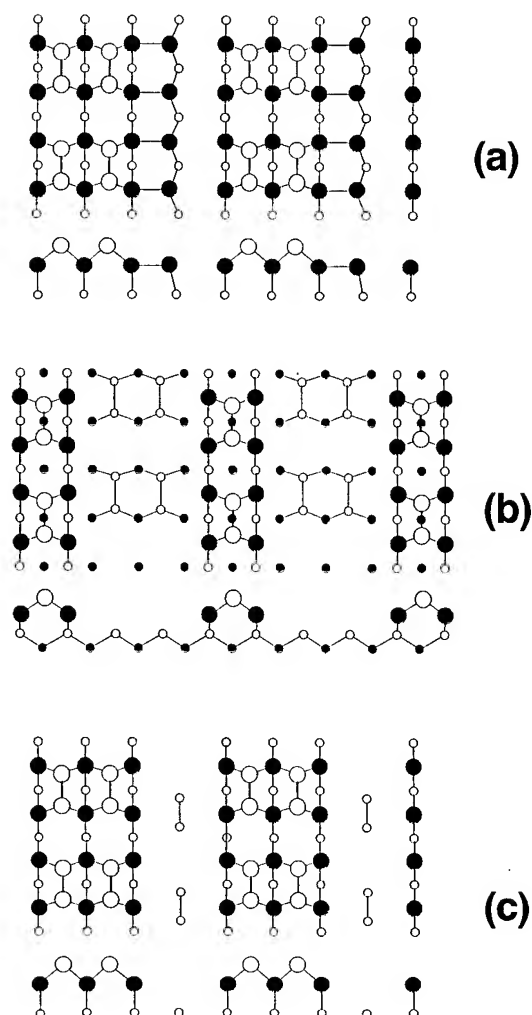


FIG. 3. The ball-to-stick models proposed for In-rich 4×2 surface. (a) Two-dimer model in this work; (b) one-dimer model by Ohkouchi and N. Ikoma (Ref. 20), and (c) bilayer two-dimer model by Biegelsen *et al.* (Ref. 24). The black solid circles correspond to the As atoms, while the open circles the In(Ga) atoms.

such problem with the so-called bilayer two-dimer model [see Fig. 3(c), Ref. 24], however, a symmetric and nodelike structure for the 1× spots should have appeared, similar to the case of Ga-rich 4×2 surface (Ref. 22). In addition, both models produce a surface In coverage of 0.75 ML. As a result, we have concluded our discussion with the two-dimer model although other possibilities may still exist.

Once the substrate is wetted with the above 4×2 structure, the progressive InAs multilayer can be grown two-dimensionally under a very simple and strict condition; the In/As₄ flux ratio, substrate temperature, and deposition rate must be chosen so that the growing front could reproduce the 4×2 symmetry. To demonstrate this, in Figs. 4(a) and 4(b), we show two STM images obtained from the surfaces deposited with 4 and 13 ML (they are much thicker than the critical thickness of ~2 ML) of InAs, respectively. No 3D island is observed at this stage and the surface is basically flat at the atomic scale, indicating a novel layer-by-layer growth.

As the use of MEE technique and extremely In-rich con-

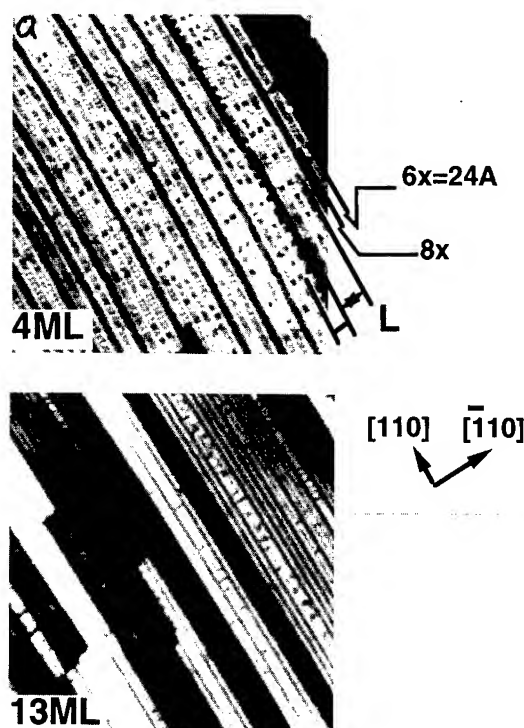


FIG. 4. The surface morphologies after (a) 4 ML (the upper panel), (b) 13 ML (the lower panel) InAs epilayer deposition on the GaAs(001) substrate under In-rich growth condition. It is evident that with increasing epilayer thickness, the formation of domain walls becomes more and more frequent.

ditions assure enough adatom diffusion, the change in the growth mode is obviously not due to limited kinetics. If it is this case, one would expect an earlier 2D–3D transition. We therefore agree with Snyder *et al.* that the layer-by-layer growth mode observed presently is due to increasing surface tension;⁹ the In-rich reconstruction has a much higher surface tension than that of the As-rich one. In traditional Stranski–Krastanov growth of a strained layer, 3D islanding relieves the elastic energy, but increases the surface area and thereby costs surface energy. The relaxation of elastic energy is dominant and 2D–3D transition is favorable where the surface tension is small. With increasing surface tension, however, the driving force for such a picture diminishes. In order that the energy gain due to a strain reduction in the islands outweighs the surface energy increase, the 2D–3D transition must be delayed. The calculation of surface tension for the In-rich 4×2 and the critical thickness for the In-rich growth condition forms an interesting topic for future theoretical investigation.

Despite the smooth surface morphology, the STM images [Figs. 4(a) and 4(b)] do not show a perfect 4×2 ordering compared to Fig. 2, although the whole RHEED displays a 4×2 symmetry. Now the surface mainly contains the 4×2 domain but is modulated with characteristic dark lines running from a step edge to another along the [110] direction, forming a unique domain wall structure with a characteristic wall separation of Na_0 (a_0 is the surface lattice constant). Statistically, $N = 6$ is dominant for both 4 and 13 ML cases. The domain wall structure appears at the same position in the

empty states image, indicating its geometric origin. These features are similar to the $2 \times N$ structure on the strained Ge/Si interface and the vacancy defect bands induced by Ni on the Si(001) surface,^{25–28} which has been attributed to be a consequences of surface strain relief. We believe that these regularly distributed “walls” may represent a new type of strain relaxation mechanism, and will be important and potential sites for the nucleation of misfit dislocations. Since no screw and other misfit dislocations were observed by the STM, their relationship is still unclear at present. A corresponding TEM experiment will be helpful to understand the question.

Recently, a similar study of the novel growth, called as virtual-surfactant mediated epitaxy, has been reported by Behrend, Wasserman, and Ploog.²⁹ The large scale STM images also reveal atomically flat morphology. Based on the high resolution images, they proposed a ball-to-stick model for the 4×2 which is different from ours. The three-dimers model fails to explain the off-phase arrangement of the 2× and 1× humps along the $[-110]$ in our image [Fig. 2(b)]. This is probably due to different preparation conditions and different phases may exist. In spite of these, the stability of all models needs to be substantiated by, for example, total energy calculations.

ACKNOWLEDGMENT

One of the authors (Q.X.) wishes to thank Dr. J. Behrend for providing their manuscript before publication.

- ¹J. Tersoff, C. Teichert, and M. G. Lagally, *Phys. Rev. Lett.* **76**, 1675 (1996).
- ²Q. Xie *et al.*, *Phys. Rev. Lett.* **75**, 2542 (1995).
- ³G. S. Solomon *et al.*, *Phys. Rev. Lett.* **76**, 952 (1996).
- ⁴G. S. Solomon *et al.*, *Appl. Phys. Lett.* **66**, 3161 (1995).
- ⁵S. Farad *et al.*, *Appl. Phys. Lett.* **65**, 1388 (1994).
- ⁶H. Drexler *et al.*, *Phys. Rev. Lett.* **73**, 2252 (1994).
- ⁷W. J. Schaffer *et al.*, *J. Vac. Sci. Technol. B* **1**, 688 (1983).
- ⁸E. Tournie *et al.*, *Appl. Phys. Lett.* **60**, 2877 (1992); E. Tournie and K. H. Ploog, *Thin Solid Films* **231**, 43 (1993).
- ⁹C. W. Snyder *et al.*, *Appl. Phys. Lett.* **62**, 46 (1993).
- ¹⁰M. Y. Yen *et al.*, *Surf. Sci.* **174**, 606 (1986).
- ¹¹V. Bressler *et al.*, *Phys. Rev. Lett.* **74**, 3209 (1995).
- ¹²C. W. Snyder *et al.*, *Phys. Rev. B* **46**, 9551 (1992).
- ¹³C. W. Snyder *et al.*, *Phys. Rev. Lett.* **66**, 3032 (1991).
- ¹⁴D. J. Eaglesham and M. Cerullo, *Phys. Rev. Lett.* **64**, 1943 (1990).
- ¹⁵T. Hashizume *et al.*, *Phys. Rev. Lett.* **73**, 2208 (1994); *Phys. Rev. B* **51**, 4200 (1995).
- ¹⁶H. H. Farrell and C. J. Palmstrom, *J. Vac. Sci. Technol. B* **8**, 903 (1990).
- ¹⁷D. J. Chadi, *J. Vac. Sci. Technol. A* **5**, 834 (1987).
- ¹⁸The deviation of growth conditions results in the normal S–K mode (with more As) or causes nonstoichiometric deposition—extra In-cluster formation (with more In).
- ¹⁹H. Yamaguchi and Y. Horokoshi, *Jpn. J. Appl. Phys.* **33**, L1423 (1994).
- ²⁰S. Ohkouchi and I. Tanaka, *Appl. Phys. Lett.* **59**, 1588 (1991).
- ²¹C. Kendrick, G. LeLay, and A. Kahn, *Phys. Rev. B* **54**, 17 877 (1996).
- ²²T. -C. Chang and D. E. Aspnes, *Phys. Rev. B* **41**, 12002 (1990); D. E. Aspnes *et al.*, *Phys. Rev. Lett.* **64**, 192 (1990); Q. K. Xue *et al.*, *ibid.* **74**, 3177 (1995).
- ²³M. D. Pashley, *Phys. Rev. B* **40**, 10481 (1990).
- ²⁴D. K. Biegelsen, R. D. Bringans, J. E. Northrup, and L. -E. Swartz, *Phys. Rev. B* **41**, 5701 (1990).
- ²⁵A. Tranpert, E. Tournie, and K. H. Ploog, *Appl. Phys. Lett.* **66**, 2265 (1995).
- ²⁶Y. W. Mo and M. G. Lagally, *J. Cryst. Growth* **111**, 876 (1991).
- ²⁷R. Butz and S. Kampers, *Appl. Phys. Lett.* **61**, 1307 (1992).
- ²⁸H. J. W. Zandvliet *et al.*, *Phys. Rev. Lett.* **75**, 3890 (1995).
- ²⁹J. Behrend, M. Wasserman, and K. H. Ploog, *J. Cryst. Growth* (in press).

Subnanometer analysis of molecular beam epitaxy grown ternary arsenides*

M. L. Seaford

Cornell University, Ithaca, New York 14853 and Wright Laboratory, Materials Directorate, Wright-Patterson Air Force Base, Ohio 45433

W. Wu

University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

K. G. Eyink and D. H. Tomich

Wright Laboratory, Materials Directorate, Wright-Patterson Air Force Base, Ohio 45433

J. R. Tucker

University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

L. F. Eastman

Cornell University, Ithaca, New York 14853

(Received 12 January 1997; accepted 5 May 1997)

InGaAs/InAlAs modulation-doped field-effect transistors (MODFETs) grown by molecular beam epitaxy on InP have been characterized using Hall and cross-sectional scanning tunneling microscopy (XSTM). Spacer layer to quantum well interface XSTM images will be presented showing cluster sizes ranging from 20 to 50 Å for normal growth conditions. The extent of the clustering experimentally found is greater than that predicted by theory for phase separation due to kinetics or bulk thermodynamics. The extent of clustering was not found to be a function of growth temperature while the morphology and cluster orientation was a function of growth temperature. At normal growth temperatures, the clusters tended to self align into binary corridors. Control (and removal) of these binary corridors allowed the growth and fabrication of MODFETs with Schottky barrier heights greater than 1 eV. © 1997 American Vacuum Society. [S0734-211X(97)11404-4]

I. INTRODUCTION

The InGaAs/InAlAs material system has demonstrated material parameters that allow for the fabrication of superior device structures such as MODFETs. However, in order to take advantage of these material properties, these devices have been fabricated with increasingly smaller dimensions and corresponding increasingly tighter control with regard to the individual layers and interfaces. With typical transistor gates, Schottky barriers, quantum wells, and spacer layer dimensions on the order of 1000, 200, 100, and 40 Å, respectively, the material layers must be analyzed and optimized at the subnanometer level. At these dimensions, the final atomic arrangement of molecular beam epitaxy (MBE) grown ternary alloys must be examined due to the various growth phenomena such as random deposition, phase separation or "clustering," and ordering.

Clustering is the degree of phase separation found in the ternary $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ (InAlAs) that exceeds the quantity that would be present due to random fluctuations.¹ For the InAlAs ternary, any composition must be constructed using the combinations of the primitive cells; InAs, $\text{In}_{0.75}\text{Al}_{0.25}\text{As}$, $\text{In}_{0.50}\text{Al}_{0.50}\text{As}$, $\text{In}_{0.75}\text{Al}_{0.25}\text{As}$, AlAs allowed by the zincblende structure. Experimental studies have shown evidence that the phase separation which occurs in these ternaries results in pure binary clusters of InAs and AlAs. When the cluster shape is approximated with a sphere, these binary

clusters have been found to have diameters of less than 10 nm. Chu *et al.*² concluded that AlAs clusters were present at certain growth conditions from observations of Schottky barrier heights at 1.2 eV in InAlAs. Using cross-sectional transmission electron microscopy (XTEM), McDevitt *et al.*³ found the average periodicity of the compositional fluctuations was on the order of 3–7 nm for a lattice matched ternary on indium phosphide (InP). Oh *et al.*⁴ modeled the results from variable temperature Hall measurements on a single InAlAs layer. Their results indicated clusters with an equivalent spherical diameter of 7 nm and required a conduction band offset of 0.6–0.7 eV within the InAlAs layer. Due to the effects of hydrostatic strain on the band gaps of various InAlAs compositions, complete phase separation of InAlAs into AlAs and InAs would be necessary to achieve this offset. Each of the previous experimental studies were performed on ternaries created using molecular beam epitaxy (MBE).

The growth processes involved during MBE have been described as a "situation [that] is usually far from equilibrium."⁵ Determining if the phenomena involved during the MBE growth processes are best described using kinetics or thermodynamics has been the goal of many experiments and papers. In agreement with Joyce *et al.*,⁶ either kinetic or thermodynamic relationships could be the correct formalisms to describe growth, depending on the growth conditions used during MBE. The group III atom surface lifetime relative to the time required to experience all the

*No proof corrections received from author prior to publication.

TABLE I. Standard SD MODFET structure.

Purpose	Composition	Thickness
Cap	$n\text{-In}_{0.53}\text{Ga}_{0.47}\text{As}$	100 Å
Schottky layer	$\text{In}_{0.52}\text{Al}_{0.48}\text{As}$	200 Å
Delta doping	-Si-	...
Spacer layer	$\text{In}_{0.52}\text{Al}_{0.48}\text{As}$	40 Å
Quantum well	$\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$	300 Å
Buffer	$\text{In}_{0.52}\text{Al}_{0.48}\text{As}$	3000 Å
Substrate	Fe:InP	...

available surface states determines which formalism is more accurate. If the surface lifetime is sufficiently small due to low substrate temperatures or high group V overpressures, the MBE growth should be a purely random deposition (kinetically controlled). If the group III atom is sufficiently mobile, all the available surface microstates can be experienced allowing for local thermodynamic equilibrium. In order to provide controllable and reproducible MBE growth of ternaries, it is necessary to determine if either kinetic or thermodynamic relationships would explain the experimentally determined binary clustering. Kinetic or random deposition should not result in clusters of AlAs or InAs primitive cells. These primitive cells should be dispersed throughout the epilayer at a density given by a standard binomial distribution. Likewise, for thermodynamic models⁷ with growth conditions at temperatures above the miscibility gap temperature (T_{MG}), the concentration of AlAs and InAs primitive cells approaches that of the random distribution as the temperature is increased. In addition, these InAs and AlAs primitive cells should be uniformly dispersed throughout the epilayer.

Various authors have used XSTM to study MBE grown compound semiconductors systems. Johnson *et al.* have shown that bulk⁸ and atomic planar doped⁹ p -type acceptors such as beryllium can be imaged and accurately counted in GaAs using XSTM images. Zheng *et al.*¹⁰ demonstrated that it was also possible to image silicon donors in GaAs. Finally, both groups have shown that alloys such as AlGaAs¹¹ and InGaAs¹² can result in a binary cluster when grown pseudomorphically strained on GaAs.

II. EXPERIMENT

A Varian Gen II solid source molecular beam epitaxy (SSMBE) reactor was used to grow InAlAs/InGaAs MODFET structures on epi-ready 3 in. iron-doped indium phosphide substrates (Fe:InP). The substrates are 2 deg mis-cut off the 001 direction. The 3 in. Fe:InP substrates were mounted on indium free molybdenum blocks. The substrates were epi-ready and no chemical cleaning was used. The substrates were baked in high vacuum at approximately 200 °C for 30 min before growth. Reflection high energy electron diffraction (RHEED) was used before growth to verify that the oxides were removed by substrate heating. The temperature during growth was controlled using optical pyrometry. The optical pyrometry was calibrated using a CI system's NTM1 noncontact temperature monitoring system. The ternary growth rates were approximately 2 Å/s. During growth,

TABLE II. Standard DD MODFET structure.

Purpose	Composition	Thickness
Cap	$n\text{-In}_{0.53}\text{Ga}_{0.47}\text{As}$	100 Å
Schottky barrier layer	$\text{In}_{0.52}\text{Al}_{0.48}\text{As}$	200 Å
Upper delta doping	-Si-	...
Spacer layer	$\text{In}_{0.52}\text{Al}_{0.48}\text{As}$	40 Å
Quantum well	$\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$	200 Å
Spacer layer	$\text{In}_{0.52}\text{Al}_{0.48}\text{As}$	40 Å
Lower Delta Doping	-Si-	...
Barrier layer	$\text{In}_{0.52}\text{Al}_{0.48}\text{As}$	150 Å
Buffer	$\text{In}_{0.52}\text{Al}_{0.48}\text{As}$	2850 Å
Substrate	Fe:InP	...

the V/III ratio was greater than 100 and the wafer was rotated at approximately 1 rotation per monolayer.

The standard structures for both single doped (SD) and double doped (DD) MODFET structures are given in Tables I and II, respectively. The compositions of the lattice matched and pseudomorphically strained layers were verified using a Philips HRXRD. Cross sectional scanning tunneling microscopy (XSTM) was done on thinned samples cleaved in vacuum. The XSTM images are of the (110) surfaces. The sample preparation and techniques have been previously described.¹³ A series of buffer layer experiments consisted of variations in the growth temperatures of the buffer layer in the standard SD MODFET. A series of Schottky barrier layer experiments consisted of altering the aluminum concentration in the Schottky barrier and spacer layer of the standard SD MODFET. A series of power DD MODFET structures were designed¹⁴ using these results. The SD and DD MODFETs were fabricated into devices as previously described.¹⁵

III. RESULTS AND DISCUSSIONS

The XSTM images of the various growth temperature regions used during the buffer layer variation are shown in Fig. 1. The phase separation can be clearly seen. The morphology and orientation of clusters is seen to be a function of growth temperature. Figure 2 contains a direct comparison of low versus high growth temperatures for the InAlAs. At the lower growth temperature, the clusters have an elongated appearance with the longer dimension perpendicular to the growth direction. At the higher growth temperature, the cluster shape is more spherical. In addition, the elongated clusters are distributed evenly across the growth front. However, the spherical clusters prefer to occur close to or directly along the growth direction from previous spherical clusters. Using the image histograms, the fractional area represented by intensities greater than and less than a fixed value were calculated. The fractional areas were approximately equal for various growth temperatures. The InAs and AlAs clusters ranged from 10% to 13% and 12% to 15%, respectively. As all the growth temperatures were greater than T_{MG} , the maximum concentration of InAs and AlAs primitive cells should have been approximately 6%. However, this 6% represents the total amount of binary primitive cells that should

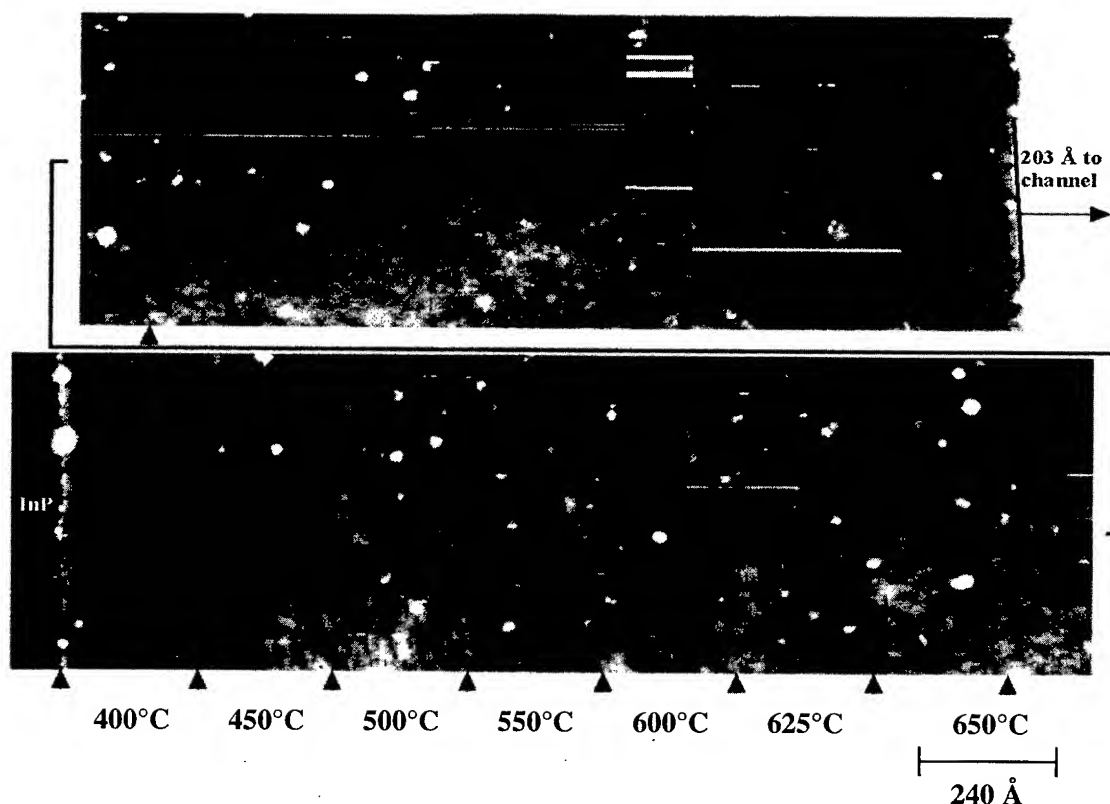


FIG. 1. Buffer layer of SD MODFET structure $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ morphology vs thermocouple temperature.

have been present. No clustering should have occurred. This result indicates that both random deposition (kinetic) and bulk thermodynamic growth models are not adequate to predict the phase separation which occurs during MBE. Bernard *et al.*¹⁶ proposed a thermodynamic model which includes surface effects that may be more adequate in explaining the phase separation that is seen during growth.

The correlated arrangement of clusters at higher growth temperatures can explain the variation in Schottky barrier height of InAlAs . At higher growth temperatures, the barrier height is independent of aluminum percentage and remains fixed at ~ 0.6 eV. At lower growth temperatures, the barrier height increases linearly with the increase in aluminum percentage. The binary clusters when aligned into a binary cor-

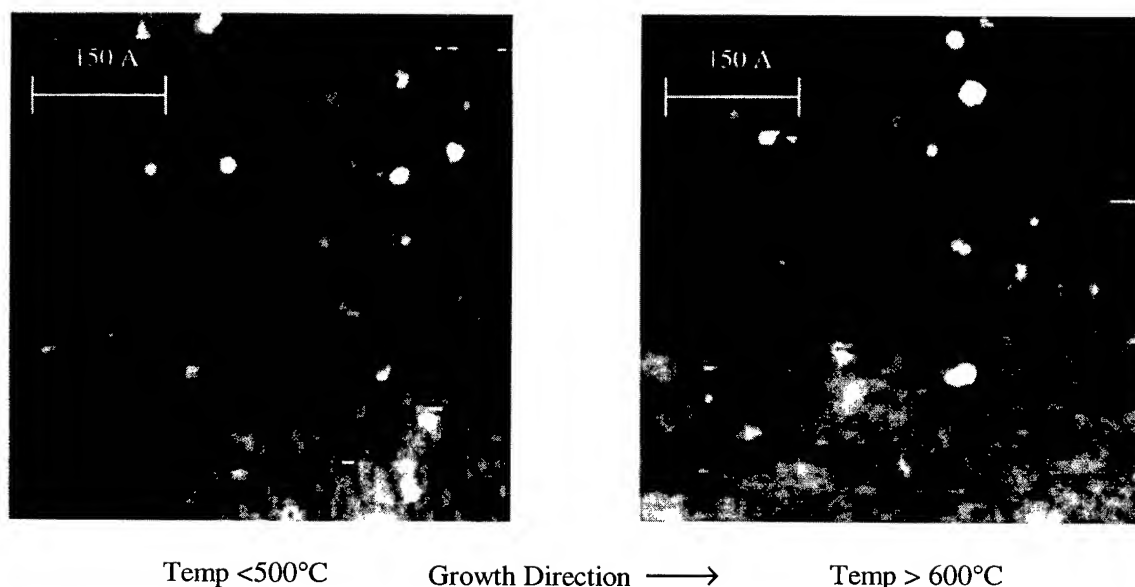


FIG. 2. Comparison of InAlAs structure due to growth temperature.



FIG. 3. Quantum well to donor spacing Si donors circled.

ridor prevent the creation of a uniform barrier and the interface which results in the lowest barrier height being measured. The use of a lower temperature barrier layer allowed the fabrication of MODFETs with Schottky barrier heights greater than 1 eV.

Figure 3 shows the InGaAs/InAlAs (quantum well/spacer) interface for the standard SD MODFET. In order to differentiate donors from adatoms, each image was taken multiple times. The resulting images when compared allowed the mobile adatoms to be distinguished. At approximately 50 Å to the right of the interface (inside the InAlAs), is the location of the delta doped layer. There appears to be at least eight donors that can be seen forming a straight line at a spacing that is reasonable for the intended donor sheet density. The quantum well/spacer interface, however, appears very rough with up to 30 Å variation in location. As can be seen by the flat growth front present during the creation of the delta

doped layer, the interface roughness of 30 Å is due to the common InAs binary cluster which is present in both the quantum well and the spacer layer. This allows for the creation of more abrupt interfaces without requiring improvements in the MBE growth process. The device structures can be designed to terminate interfaces with the desired binaries that are components within each individual layer that will enhance the overall device performance.

IV. CONCLUSIONS

The correlated use of XSTM images with variations in the growth temperature used during MBE has allowed improved device performance by design. Phase separation into binary clusters was found in InAlAs and InGaAs layers grown by solid source MBE on InP substrates. The kinetic models which rely on random deposition do not predict the observed clustering. The thermodynamic models do not predict clusters at the growth temperatures. Neither random deposition (kinetic) nor bulk thermodynamic models were capable of predicting the observed clustering. The extent of the clustering was not found to be a function of growth temperature in the range investigated. The morphology and orientation of the clusters formed in InAlAs was found to depend on the growth temperature. Realization of this control allowed the growth of DD MODFET structures with Schottky barrier layers with heights greater than 1 eV. In addition, the growth processes involved in current MBE techniques are seen to generate a smooth growth front from the nearly linear delta doped layer. The interface roughness which is due to the common binary of two adjacent ternaries can be eliminated by designing binary interface layers that will improve device performance.

V. FUTURE WORK

The correlated use of XSTM with variations in growth parameters has already demonstrated the ability to explain structure phenomena and improve device performance. However, only variations of the growth temperature has currently been investigated. Growth parameters such as growth rates, V/III ratios, substrate orientation, etc. still need to be studied. To accurately predict the microstructure of MBE grown layers, an alternate model needs to be developed, perhaps relying on surface thermodynamics.

¹A. Zunger and S. Mahajan, Handbook on Semiconductors, Vol. 3b, 1994 (unpublished), p. 1417.

²P. Chu, C. L. Lin, and H. H. Wieder, Appl. Phys. Lett. **53**, 2423 (1988).

³T. L. McDevitt, S. Mahajan, D. E. Laughlin, W. A. Bonner, and V. G. Keramidas, Phys. Rev. B **45**, 6614 (1992).

⁴J. E. Oh, P. K. Bhattacharya, Y. C. Chen, O. Aina, and M. Mattingly, J. Electron. Mater. **19**, 435 (1990).

⁵J. W. Matthews, *Epitaxial Growth Part A* (Academic, New York, 1975), p. 38.

⁶B. A. Joyce, Handbook on Semiconductors, edited by S. Mahajan, Vol. 3A, 1994 (unpublished), p. 344.

⁷A. Chen and A. Sher, *Semiconductors Alloys: Physics and Materials Engineering* (Plenum, New York, 1995).

⁸M. B. Johnson, O. Albrektsen, R. M. Feenstra, and H. W. M. Salemink, Appl. Phys. Lett. **63**, 2923 (1993).

- ⁹M. B. Johnson, P. M. Koenraad, W. C. van der Vleuten, H. W. M. Salemink, and J. H. Wolter, *Phys. Rev. Lett.* **75**, 1606 (1995).
- ¹⁰J. F. Zheng, X. Liu, N. Newman, E. R. Weber, D. F. Ogletree, and M. Salmeron, *Phys. Rev. Lett.* **72**, 1490 (1994).
- ¹¹M. B. Johnson, U. Maier, H.-P. Meier, and H. M. K. Salemink, *Appl. Phys. Lett.* **63**, 1273 (1993).
- ¹²J. F. Zheng, J. D. Walker, M. B. Salmeron, and E. R. Weber, *Phys. Rev. Lett.* **72**, 2414 (1994).
- ¹³W. Wu, Ph.D. thesis, University of Illinois, 1996.
- ¹⁴G. H. Martin, M. L. Seaford, R. Spencer, J. Braunstein, and L. F. Eastman, *Proc. of the IEEE/Cornell Conference*, Aug. 1995 (unpublished), pp. 132-140.
- ¹⁵G. H. Martin, A. Lepore, M. L. Seaford, B. Periaslavets, R. Spencer, and L. F. Eastman, *IEDM Proc.*, 1996.
- ¹⁶J. E. Barnard, S. Froyen, and A. Zunger, *Phys. Rev. B* **44**, 11 178 (1991).

Local interface composition and extended defect density in ZnSe/GaAs(001) and ZnSe/In_{0.04}Ga_{0.96}As(001) heterojunctions

S. Heun, J. J. Paggel,^{a)} L. Sorba,^{b)} S. Rubini, and A. Franciosi^{c)}
*Laboratorio Nazionale TASC-INFM, Area di Ricerca, Padriciano 99, I-34012 Trieste, Italy,
and Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis,
Minnesota 55455*

J.-M. Bonard^{d)} and J.-D. Ganière
*Institut de Micro- et Optoelectronique, Département de Physique, Ecole Polytechnique Fédérale,
CH-1015 Lausanne, Switzerland*

(Received 13 January 1997; accepted 22 April 1997)

We have recently shown that in II–VI/III–V heterojunctions and related devices fabricated by molecular beam epitaxy, the II–VI flux ratio employed during the early stages of II–VI growth can be used to control the local interface composition and the band alignment. Here we demonstrate that the local interface composition in pseudomorphic, strained ZnSe/GaAs(001) heterostructures as well as lattice-matched ZnSe/In_{0.04}Ga_{0.96}As(001) heterostructures also have a dramatic effect on the nucleation of native stacking fault defects. Such extended defects have been associated with the early degradation of blue-green lasers. We found, in particular, that Se-rich interfaces consistently exhibited a density of Shockley stacking fault pairs below our detection limit and three to four orders of magnitude lower than those encountered at interfaces fabricated in Zn-rich conditions.

© 1997 American Vacuum Society. [S0734-211X(97)09104-X]

I. INTRODUCTION

Although since 1991 substantial progress has been made in the implementation of blue-green solid-state laser diodes based on wide-gap II–VI semiconductor active layers fabricated by molecular beam epitaxy (MBE) on III–V substrates,¹ the limited lifetime of such devices is still hindering the development of a viable technology. Recent progress in eliminating dislocations from the laser structure through the use of quaternary cladding and waveguiding layers lattice matched to the III–V substrates has reduced but not eliminated the problem,¹ so that attention has been focused lately on other types of native defects which are preferentially located at the II–VI/III–V interface and may contribute to device degradation.

In particular, degradation during cw operation of pseudomorphic laser structures at room temperature has been associated in part^{2,3} with stacking faults at the ZnSe/GaAs or ZnSe_{1-y}S_y/GaAs interface. Such defects would propagate through the structure during growth and generate dislocation sources within the strained Zn_{1-x}Cd_xSe active layers.^{2,3} The existence of relatively high densities of native stacking faults at the ZnSe/GaAs interface has been known for some time⁴ and tentatively explained by mechanisms as different as island coalescence during II–VI epilayer growth,⁵ doping,⁶ or substrate roughness.⁷ Both intrinsic and extrinsic stacking faults, which are formed by removing or inserting, respec-

tively, a portion of a closed-packed {111} plane have been reported,⁷ in pairs or as isolated defects, starting in all cases at or near the II–VI/III–V interface and propagating throughout the II–VI epilayer,²⁻⁸ bounded by Frank^{2-4,7} or Shockley^{7,8} partial dislocations.

Several authors have attempted to reduce or eliminate the native stacking fault concentration by modifying the growth conditions during fabrication of the II–VI/III–V heterostructure. Gaines *et al.*⁹ using migration-enhanced epitaxy for ZnSe/GaAs(001) reported a high density (10^8 cm⁻²) of stacking faults if the III–V surface was exposed initially to a flux of elemental Se at 500 °C prior to ZnSe growth, while exposure to a Zn flux prior to II–VI growth reportedly led to much lower stacking fault densities. Similar Zn- and Se-predosing techniques were employed by Kuo *et al.*⁷ for ZnSe_{1-y}S_y/GaAs heterostructures grown by conventional MBE. Densities of Frank and Shockley dislocations at the 10^8 cm⁻² level were reported for growth on As-stabilized surfaces initially exposed to a Se flux, while densities below the experimental sensitivity ($<5 \times 10^4$ cm⁻²) were observed in samples where Zn predosing was employed. Kuo *et al.*⁷ associated the higher density of defects with the nucleation of three-dimensional (3D) islands in the early stages of growth of the II–VI epilayer on Se-exposed surfaces, mostly based on the corresponding degraded quality of the reflection high energy electron diffraction (RHEED) pattern. Similar proposals were put forth by Gaines *et al.*⁹ and Guha *et al.*¹⁰ for ZnSe/GaAs.

However, in a recent study of pseudomorphic ZnSe/GaAs(001) heterostructures grown by metal organic chemical vapor deposition (MOCVD), Bourret-Courchesne¹¹ reported an increase in the stacking fault density for heterostructures fabricated after exposure to Zn of the As-

^{a)}Present address: Philipps-Universität Marburg, Fachbereich Physik, Renthof 6, D-35032 Marburg, Germany.

^{b)}Also with Istituto ICMAT del CNR, Monterotondo, I-00016 Rome, Italy.

^{c)}Also with Dipartimento di Fisica, Università di Trieste, I-34127 Trieste, Italy.

^{d)}Present address: Institut de Physique Expérimentale, Département de Physique, Ecole Polytechnique Fédérale, CH-1015 Lausanne, Switzerland.

stabilized substrate, as well as for growth on surfaces on which excess As was present, and tentatively associated the increase with the formation of small clusters of zinc arsenide.

We focus here on the apparent discrepancy between the MOCVD and MBE results and present evidence that the Zn/Se flux ratio employed during the early stages of ZnSe growth by MBE on GaAs(001)2×4 as well as In_{0.04}Ga_{0.96}As(001)2×4 surfaces has a profound effect on the resulting density of native stacking faults. We found that the use of low Zn/Se beam pressure ratios (BPR), i.e., Se-rich growth conditions, during interface fabrication yields stacking fault densities dramatically lower than those obtained with high Zn/Se BPRs. Since we have shown earlier that the BPR controls the interface composition and the band alignment,^{12–14} the present results indicate that the local atomic configuration of the interface contributes to determine the probability of stacking fault nucleation.

In a recent article, Kuo *et al.* performed a similar experiment.¹⁵ However, prior to the ZnSe growth with different BPRs, they exposed the GaAs(001)2×4 surface to a flux of Zn. While for the Zn-rich growth conditions their results are in good agreement with our data, they obtained for BPR=0.1 a Shockley partial dislocation density of $5 \times 10^6 \text{ cm}^{-2}$. This value is clearly higher than the value reported here, probably due to the Zn predosing of the GaAs substrate.

II. EXPERIMENTAL DETAILS

All heterostructures were fabricated by MBE following the methods described elsewhere.^{12,13} The substrates were GaAs(001) or In_{0.01}Ga_{0.99}As(001) wafers, on which 1- μm -thick GaAs(001)2×4 or 1.5- to 2- μm -thick In_{0.04}Ga_{0.96}As(001)2×4 buffer layers were grown at 580 and 490 °C, respectively. ZnSe overlayers 100- and 300-nm-thick were deposited at 290 °C on the III–V buffers. In view of the 0.27% in-plane lattice mismatch for ZnSe–GaAs(001), the 100-nm-thick overlayer was pseudomorphic and the 300-nm-thick layers partially relaxed. ZnSe–In_{0.04}Ga_{0.96}As(001) heterostructures are expected to be lattice matched instead.

During II–VI growth, we used BPRs in the 0.1–10 range, as determined from an ion gauge positioned at the sample location. The BPR is known to affect *per se* the optical and structural quality of the bulk II–VI epilayer. We therefore confined nonstoichiometric growth to a thin (2-nm-thick) composition-control interface layer (CIL), and grew the rest of the II–VI epilayer with a constant BPR=1 in order to isolate the effect of the BPR on stacking fault nucleation at the II–VI/III–V interface.¹⁶ The local composition of the CILs was calibrated *in situ* using a monochromatic x-ray photoemission spectrometer (XPS) through the integrated intensity of the Zn 3d, Se 3d, Ga 3d, and As 3d core doublets. We have shown earlier^{12–14} that large Zn overpressures (high BPRs) during the early stages of growth yield Zn-rich interface compositions, while large Se overpressures (low BPRs) correspond to Se-rich interface compositions. The XPS results would be consistent with excess Zn (or Se) elemental concentrations at the monolayer level of the interface, below

an essentially stoichiometric II–VI overlayer. The local interface composition was found to control, in turn, the band offsets and the electrical properties of the heterojunction.^{12–14}

The extended defect density for each sample was determined using transmission electron microscopy (TEM). Plan views and cross sections were obtained by ion milling the sample at room temperature using 5 keV Ar⁺ ions at an angle of incidence of 4°. A Philips EM430ST TEM operated at 300 keV allowed us to sample the overall heterostructure in sections as thick as 2 μm . The observation of thin regions of the plan-view thin foils allowed us to focus on the structural quality of the II–VI epilayer alone, while thicker regions yielded information on the whole sequence of epilayers down to the III–V buffer/substrate interface. The detection limit for stacking faults in our experiment was $1 \times 10^5 \text{ cm}^{-2}$. A single image with a magnification of 10 000× represents a sampled area of $8 \times 10^{-7} \text{ cm}^2$. A detection limit of $1 \times 10^5 \text{ cm}^{-2}$ means therefore that no stacking fault was observed in 13 images from different spots of the same sample.

III. RESULTS AND DISCUSSION

The dominant type of stacking faults observed in both ZnSe–GaAs(001) and ZnSe–In_{0.04}Ga_{0.96}As(001) heterostructures is exemplified in the plan-view, bright-field TEM images in Fig. 1. We show images obtained from a 100-nm- and 300-nm-thick ZnSe overlayer grown on In_{0.04}Ga_{0.96}As(001). The photographs in the left and right columns were obtained using (000)-(220) and (000)-(220) two-beam diffraction conditions, respectively, near the [001] zone axis. Within the 100-nm-thick overlayers, by far the most numerous defects are stacking fault pairs, which originate at the II–VI/III–V interface, that extend throughout the overlayer along the (111) and $(\bar{1}\bar{1}\bar{1})$ planes, and intersect the growth surface along the [110] direction. Some isolated stacking faults are also visible, oriented perpendicular to the pairs, and exhibit well-defined contrast fringes.

The stacking fault pairs and the isolated stacking faults in Fig. 1 can be identified as Shockley and Frank stacking faults, respectively, based on the characteristic contrast they exhibit in the different diffraction conditions.^{17,18} First, near the [001] axis for the (220) reflection, the partial dislocations bounding Frank stacking faults should become invisible, which is not the case for the dislocations bounding the stacking fault pairs in Fig. 1, but is instead verified for the isolated faults in Fig. 1. Second, for each stacking fault pair, images recorded near the [114] axis with the $(\bar{1}31)$ and $(3\bar{1}1)$ reflections (not shown), clearly depict the (111) fault of each pair while the $(\bar{1}\bar{1}\bar{1})$ fault vanishes. Third, the two partial dislocations bounding the (111) fault of each pair and pointing in the [010] and [100] directions disappear in images recorded near the [114] axis with $(\bar{1}31)$ and $(3\bar{1}1)$ reflections, respectively. The isolated Frank stacking faults are instead completely unaffected. These results imply that the (111) and $(\bar{1}\bar{1}\bar{1})$ faults of each pair are of the Shockley type, with partial dislocations exhibiting the Burgers vectors $(1/6)[\bar{1}2\bar{1}]$

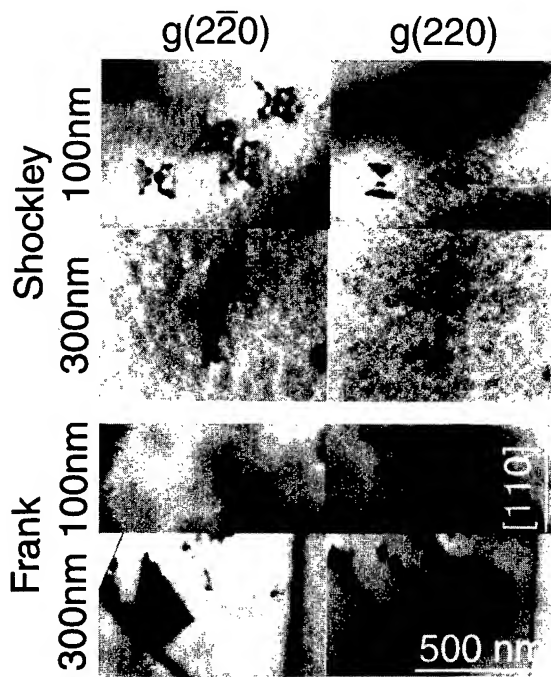


FIG. 1. Plan-view, bright-field transmission electron microscopy (TEM) images of typical Shockley- (top) and Frank- (bottom) type stacking faults observed in ZnSe-In_{0.04}Ga_{0.96}As(001) heterostructures. Images are shown for both 100-nm- and 300-nm-thick ZnSe overlayers, scaled to the same magnification and oriented with the [110] direction pointing to the top. The photographs in the left and right column were obtained using (000)-(220) and (000)-(220) two-beam diffraction conditions, respectively, near the [001] zone axis. The Shockley stacking faults always appeared in pairs and changed their shape with increasing layer thicknesses. The Frank stacking faults retained their shape and are oriented perpendicular to the Shockley pairs.

and $(1/6)[\bar{2}11]$ for the former, and $(1/6)[\bar{2}1\bar{1}]$ and $(1/6)[\bar{1}21]$ for the latter.^{17,18}

Other peculiar differences between Shockley and Frank stacking faults become evident with increasing overlayer thickness. For a ZnSe thickness of 300 nm, the Shockley pairs assume the elongated profile shown in the second row of Fig. 1, and already have been reported by a number of authors.^{7,8} While the stacking fault pairs maintain the characteristic contrast variations with varying diffraction conditions described in the previous paragraph, the change in their appearance may reflect an elongation and distortion of the lines of the partial dislocations bounding the faults. Shockley partial dislocations have Burgers vectors parallel to the glide plane, and are in principle comparatively mobile. The partial dislocations bounding the Frank stacking faults can instead only move by climbing and correspondingly the Frank stacking faults are seen to retain their characteristic appearance at all overlayer thicknesses.

The Zn/Se BPR employed during the early stages of interface formation has a dramatic effect on the density of stacking faults. This is shown in Fig. 2 for pseudomorphic ZnSe/GaAs(001) heterostructures comprised of 100-nm-thick overlayers and incorporating different CILs. The TEM plan view, bright-field micrographs were recorded in the (000)-(220) two-beam diffraction condition, which empha-

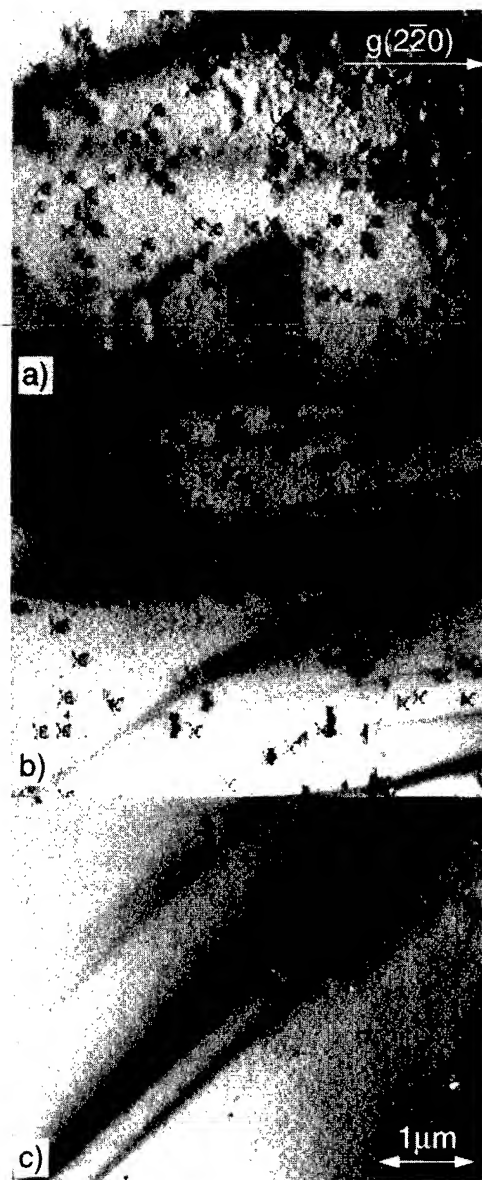


FIG. 2. TEM plan-view, bright-field micrographs of pseudomorphic ZnSe/GaAs(001) heterostructures comprised of 100-nm-thick overlayers. (a) Sample incorporating a Zn-rich composition-control interface layer (CIL) grown with a Zn/Se beam pressure ratio (BPR) of 10. (b) Control sample grown with a BPR=1 throughout. (c) Sample incorporating a Se-rich CIL grown with a BPR=0.1. The (000)-(220) two-beam diffraction condition emphasizes the Shockley partial dislocations that bound the stacking fault pairs. The areal density of Shockley stacking faults pairs varies from $7 \times 10^8 \text{ cm}^{-2}$ in (a) to $1 \times 10^5 \text{ cm}^{-2}$ in (c).

size the Shockley partial dislocations that bound the stacking fault pairs.^{17,18} Figure 2(a) depicts a heterostructure which incorporates a Zn-rich CIL (grown with a BPR=10), and exhibits an areal density of Shockley stacking faults of $7 \times 10^8 \text{ cm}^{-2}$. The control sample in Fig. 2(b) was grown with a BPR=1 throughout, and shows a density of Shockley stacking faults 3 to 4 times lower, i.e., on the order of $2 \times 10^8 \text{ cm}^{-2}$. The heterostructure depicted in Fig. 2(c) incorporated a Se-rich CIL (grown with a BPR=0.1), and exhibits

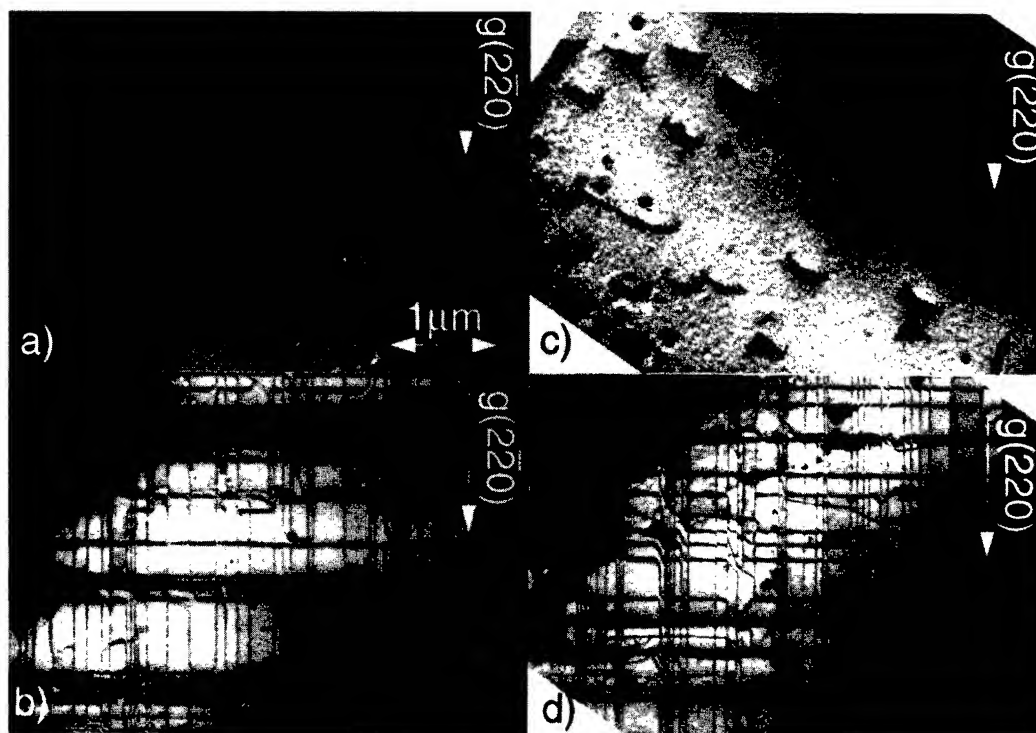


FIG. 3. TEM plan-view, bright-field micrographs of ZnSe-In_{0.04}Ga_{0.96}As(001) heterostructures incorporating a Se-rich CIL grown with a BPR=0.1 [(a) and (b)] or a Zn-rich CIL grown with a BPR=10 [(c) and (d)]. All In_{0.04}Ga_{0.96}As buffers were fabricated on GaAs(001) wafers and included a 300-nm-thick II-VI overlayer. We show images obtained from thick [(b) and (d), ~2-μm-thick] and thin [(a) and (c), ~500-nm-thick] regions of the thin foil and recorded with $g=(2\bar{2}0)$ two-beam diffraction conditions. The misfit dislocation network at the In_{0.04}Ga_{0.96}As/GaAs interface [see (b) and (d)] has a negligible effect on the structural quality of the II-VI epilayers. The areal density of Shockley stacking fault pairs varies from $1.6 \times 10^8 \text{ cm}^{-2}$ in (c) to $< 2 \times 10^5 \text{ cm}^{-2}$ in (a).

a much lower Shockley stacking fault density of $1 \times 10^5 \text{ cm}^{-2}$.

The main result in Fig. 2, i.e., that samples incorporating a Se-rich interface layer exhibit an areal density of Shockley stacking fault pairs lower by three to four orders of magnitude as compared to the others, is more difficult to verify in ZnSe/GaAs(001) with a 300-nm-thick overlayers, because of the dominant contrast associated with the formation of the large density of misfit dislocations.

ZnSe-In_{0.04}Ga_{0.96}As(001) heterostructures represent ideal systems to assess the role of strain and interface composition in determining the native stacking fault density in II-VI/III-V heterostructures. Because of the ideal lattice match, higher overlayer thicknesses can be fabricated without giving rise to the formation of misfit dislocations, and the local interface composition can still be controlled via the CIL technique. In Fig. 3, we compare bright field, plan view micrographs of ZnSe-In_{0.04}Ga_{0.96}As(001) heterostructures incorporating a 300-nm-thick II-VI overlayer and a Se- (left) or Zn-rich (right) CIL. Micrographs (a) and (c) were obtained from thin regions of the thin-foil sample and illustrate the structural quality of the II-VI epilayers. Micrographs (b) and (d) were obtained from thicker regions of the thin-foil sample, and emphasize the dislocation network present deeper in the heterostructure.¹⁹

The misfit dislocation network—due to strain relaxation of the In_{0.04}Ga_{0.96}As buffer on GaAs—is seen to have a neg-

ligible effect on the structural quality of the II-VI epilayers. TEM cross-sectional results (not shown) confirm that the dislocation network is localized at the In_{0.04}Ga_{0.96}As/GaAs interface and that there is a negligible penetration of threading dislocations from the III-V buffer into the II-VI overlayer within our experimental uncertainty.¹⁸ The only defects visible in the II-VI overlayers are stacking faults in Fig. 3(c) which originate from the II-VI/III-V interface. The $g=(2\bar{2}0)$ two-beam diffraction conditions employed in Fig. 3 emphasize the partial dislocations that bound the stacking fault pairs, which have the characteristically elongated shape already observed in Fig. 1 in this overlayer thickness range. While the heterostructure incorporating the Zn-rich CIL in Fig. 3(c) shows an areal density of Shockley stacking faults on the order of $1.5 \times 10^8 \text{ cm}^{-2}$, the density of Shockley faults in the sample incorporating the Se-rich CIL in Fig. 3(a) is below our TEM experimental sensitivity of $2 \times 10^5 \text{ cm}^{-2}$.

A more quantitative comparison of the results for ZnSe/GaAs(001) and ZnSe/In_{0.04}Ga_{0.96}As(001) heterostructures is shown in Table I, where we list the areal density of Shockley stacking fault pairs (column 4), Frank stacking faults (column 5), and threading dislocations (column 6), for the different samples (columns 1–3). The data in Table I clearly indicates that the parameter that most strongly influences the concentration of native stacking faults is the composition of the interface region, as opposed to overlayer thickness or

TABLE I. Quantitative comparison of the areal density of Shockley stacking fault pairs (column 4), Frank stacking faults (column 5), and threading dislocations (column 6), for different ZnSe/GaAs(001) and ZnSe/In_{0.04}Ga_{0.96}As(001) heterostructures (column 1) with ZnSe overlayer thickness of 100 or 300 nm (column 2), and incorporating a 2-nm-thick CIL grown with a Zn/Se beam pressure ratio (BPR) of 10 or 0.1, or fabricated with a BPR=1 throughout (column 3). The data were obtained from series of TEM plan view, bright-field micrographs. The experimental sensitivity was $2 \times 10^5 \text{ cm}^{-2}$ unless noted otherwise. Different markers (*, +) in column 1 denote In_{0.04}Ga_{0.96}As buffers grown on GaAs(001) and In_{0.01}Ga_{0.99}As(001) wafers, respectively.

Sample	ZnSe (nm)	CIL (BPR)	Shockley (cm^{-2})	Frank (cm^{-2})	Threading dislocations (cm^{-2})
ZnSe/GaAs, No. 323	100	10	7.0×10^8	3.0×10^6	$< 2 \times 10^5$
ZnSe/GaAs, No. 499	100	10	5.6×10^8	2.3×10^7	$\sim 6 \times 10^7$
ZnSe/InGaAs*, No. 500	100	10	9.2×10^8	3.2×10^7	$< 2 \times 10^5$
ZnSe/InGaAs ⁺ , No. 510	100	10	7.0×10^8	1.8×10^6	$< 2 \times 10^5$
ZnSe/InGaAs*, No. 502	300	10	1.6×10^8	1.8×10^7	$\sim 1 \times 10^7$
ZnSe/InGaAs ⁺ , No. 501	300	10	1.5×10^8	7.0×10^6	$< 2 \times 10^5$
ZnSe/GaAs No. 326	100	1	2.0×10^8	5.0×10^7	$< 2 \times 10^5$
ZnSe/GaAs No. 488	100	1	2.6×10^6	7.5×10^5	$\sim 2 \times 10^6$
ZnSe/InGaAs*, No. 511	100	1	6.8×10^6	2.6×10^6	$< 2 \times 10^5$
ZnSe/InGaAs*, No. 513	300	1	5.2×10^6	1.7×10^6	$< 2 \times 10^5$
ZnSe/InGaAs ⁺ , No. 514	300	1	1.0×10^7	2.6×10^6	$\sim 3 \times 10^6$
ZnSe/GaAs, No. 330	100	0.1	1×10^5	2.0×10^6	$< 2 \times 10^5$
ZnSe/GaAs No. 504	100	0.1	$< 5 \times 10^4$	2.4×10^6	$< 5 \times 10^4$
ZnSe/InGaAs*, No. 506	100	0.1	$< 2 \times 10^5$	6.1×10^6	$< 2 \times 10^5$
ZnSe/InGaAs ⁺ , No. 505	100	0.1	$< 2 \times 10^5$	$< 2 \times 10^5$	$< 2 \times 10^5$
ZnSe/InGaAs*, No. 509	300	0.1	$< 2 \times 10^5$	1.7×10^6	1.7×10^6
ZnSe/InGaAs ⁺ , No. 508	300	0.1	$< 2 \times 10^5$	2.2×10^6	$< 2 \times 10^5$

strain. The vast majority of the stacking faults are Shockley stacking fault pairs, but their number decreases by at least three to four orders of magnitude in going from Zn- to Se-rich interfaces. The minimum stacking fault densities observed in Table I compare favorably with the best results reported in the literature.⁴⁻¹²

The density of single isolated Frank stacking faults is also affected by the interface composition, although to a lesser extent. A decrease in the density of Frank stacking faults by a factor of 5-10 is typically observed in Table I when comparing heterostructures with Zn- and Se-rich interfaces. The implication is that the microscopic mechanism behind the formation of Shockley and Frank stacking faults at the II-VI/III-V interface might be qualitatively different, although both are influenced by the Zn/Se atomic flux ratio employed in the early stage of interface formation.

Our results seem to be in sharp contrast with the reported effect of Zn and Se predosing during MBE fabrication, since Zn predosing has been associated with enhanced two-dimensional (2D) growth and lower stacking fault densities, and Se predosing with three-dimensional (3D) growth and higher stacking fault densities.^{7,9,10} We propose therefore that the mechanism through which the local interface composition affects the stacking fault nucleation rate is qualitatively different from that proposed in Refs. 7, 9, and 10 to explain the effect of Se or Zn predosing. This is qualitatively demonstrated by a comparison of the RHEED patterns observed during growth of Se- and Zn-rich CILs in Fig. 4. Sharp 2×1 and $c(2 \times 2)$ patterns are observed in the two cases, with no indication of 3D growth in the Se-rich case.

Degraded and spotty RHEED patterns are instead observed during growth on Se-predosed surfaces.^{6,7}

We speculate that the apparent discrepancy might reflect the different Zn or Se-coverage obtained on the GaAs surface with the two methods. ZnSe deposition in Zn- or Se-rich conditions during CIL fabrication yields symmetric changes

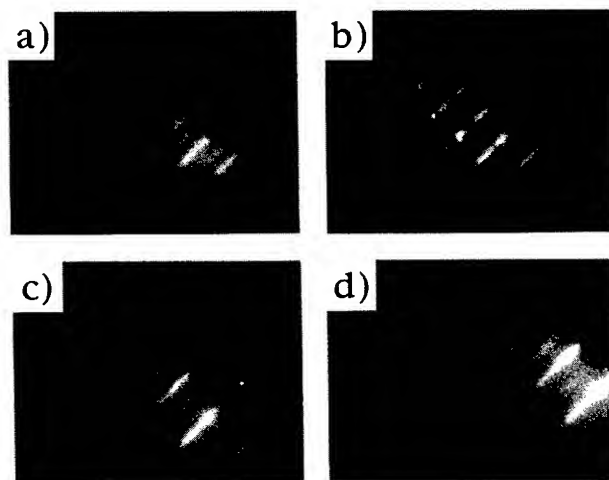


FIG. 4. Reflection high energy electron diffraction (RHEED) patterns observed during the fabrication of CILs with a BPR=0.1 (top) and a BPR = 10 (bottom). The sample azimuth was $\langle 110 \rangle$ in (a) and (b), and $\langle 100 \rangle$ in (c) and (d). The primary electron energy was 10 keV, the ZnSe film thickness for these photographs was 2 nm. Equally sharp 2×1 and $c(2 \times 2)$ RHEED patterns are observed in the two cases, with no evidence of three-dimensional growth.

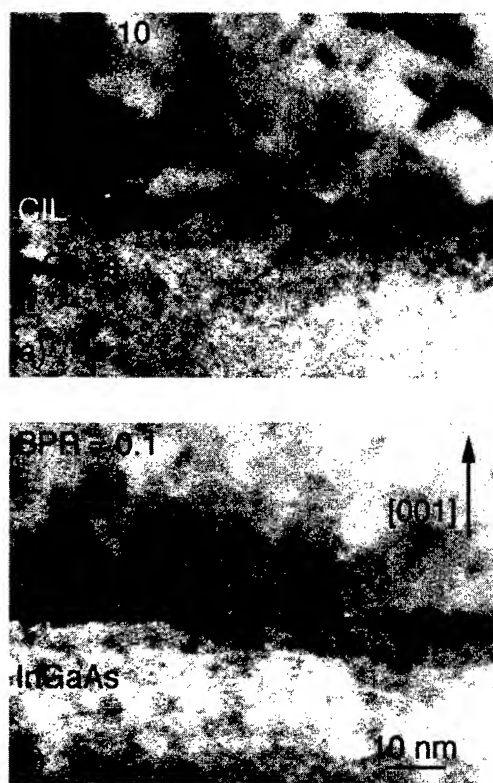


FIG. 5. Bright field cross-sectional images of ZnSe/In_{0.04}Ga_{0.96}As(001) heterostructures incorporating a Zn- (a) or Se-rich (b) CIL, grown with a BPR of 10 or 0.1, respectively. Both interfaces are viewed along the [100] zone axis, but observations along the [110] zone axis yielded similar results.

in the local interface composition on the monolayer scale, as demonstrated by our *in situ* core level photoemission studies.²¹ On the other hand, predosing at typical MBE temperatures for relatively short times²⁰ may yield lower Zn and Se coverages. For example, only Zn saturation coverages well below one monolayer can be achieved during Zn predosing,²¹ so that while Zn surface impurities may still affect step and kink energetics, and adatom mobilities, they are unlikely to result in a substantial change in the interface composition.

The trends in Table I clearly suggest that the local interface composition determines the probability of stacking fault nucleation. This might occur directly through the BPR-induced difference in the local atomic configuration of the surface/interface, or indirectly as a result of the different stability or lateral homogeneity of the surface/interface. Heterostructures fabricated in Se-rich conditions appear less stable than those obtained in Zn-rich conditions, since annealing at relatively low temperatures (300–350 °C) gives rise to a strong enhancement in the deep level emission from Zn vacancies, or complex defect centers involving Zn vacancies and substitutional Ga atoms on Zn sites.²² This might imply that such interfaces rearrange themselves more readily during growth as compared to Zn-rich configurations.

As far as the lateral homogeneity is concerned, TEM indicates important differences for Se- and Zn-rich interfaces. Figure 5 shows bright field cross-sectional images of

ZnSe/In_{0.04}Ga_{0.96}As(001) heterostructures incorporating a Zn- (a) and Se-rich (b) CIL. Both interfaces are viewed along the [100] zone axis, but observations along the [110] zone axis yielded similar results. The Zn-rich sample (a) shows a dark contrast forming a dotted line along the interface, with a dot width of $\sim 0.7 \pm 0.2$ nm about the interface and a typical length along the interface of 2 nm. For the Se-rich sample (b), we observe a dark contrast giving rise to a continuous, $\sim 0.7 \pm 0.2$ -nm-wide layer at the interface. The contrast in Fig. 5 is not due to strain but presumably reflects atomic interdiffusion and/or the formation of some type of interface reaction product. Observations in high resolution did not offer further indications as to the origin of the contrast.

Recent synchrotron radiation photoemission studies of ZnSe/GaAs(001) interfaces have revealed enhanced atomic interdiffusion across Se-rich interfaces as compared to Zn-rich interfaces.²³ Earlier studies²⁴ have associated interfacial contrasts in TEM images with the presence of an interfacial compound, Ga₂Se₃, under certain growth conditions (typically Ga-rich substrate surfaces). Ga₂Se₃ has a sphalerite structure where a third of the Ga sites are vacant, thereby modifying the structure factor of the crystal as compared to ZnSe and GaAs and causing a strong contrast in TEM. The photoemission signature of Ga₂Se₃, however, is absent from interfaces fabricated on As-rich surfaces.²³ Alternatively, other authors have recently proposed the formation of a ZnAs layer with a zincblende structure at the interface.²⁵

One could speculate that, whatever the character of the interface reaction product in Fig. 5 might be, the lateral inhomogeneities observed in the Zn-rich case may promote stacking fault formation. For example, the inhomogeneous interface reaction in the Zn-rich case may yield the nucleation of threading defects at the boundaries between bright and dark interface regions. These defects would form, in turn, stacking fault pairs. Shockley partial dislocations are formed by atomic displacements as a result of shear stress, and could therefore arise because of the strain and stacking errors resulting from the lateral inhomogeneities. On the other hand, Frank stacking faults are induced by stacking errors during the growth, and their density is therefore presumably not strongly influenced by the BPR technique.

IV. CONCLUSIONS

We identified Shockley stacking faults pairs as the dominant type of stacking fault defects in ZnSe/GaAs(001) and ZnSe/In_{0.04}Ga_{0.96}As(001) heterostructures, and found that fabrication of Se-rich interfaces results in a decrease by three to four orders of magnitude in the density of such stacking faults as compared to Zn-rich interfaces. We observed qualitatively similar 2D growth during fabrication of Se- and Zn-rich interfaces, so that the influence of the BPR on the stacking fault density must reflect a mechanism qualitatively different from that proposed earlier by other authors to explain the effect of Zn and Se predosing of the III–V surface prior to II–VI growth.

Since similar trends are observed for both strained pseudomorphic ZnSe/GaAs heterostructures as well as

lattice-matched ZnSe/In_{0.04}Ga_{0.96}As(001) heterostructures, strain does not appear to play a major role in determining the stacking fault density.

The different atomic termination, thermodynamic stability, and lateral homogeneity of Se- and Zn-rich interfaces might all, in principle, play a role in determining the widely different probability of stacking fault nucleation in these heterojunctions.

ACKNOWLEDGMENTS

The work in Trieste was funded in part by the Consiglio Nazionale delle Ricerche under the NOVA project. The work in Minneapolis was supported in part by the U.S. Army Research Office under Grant Nos. DAAH04-93-G-0206 and DAAH04-93-G-0319, and by the Center for Interfacial Engineering of the University of Minnesota. We acknowledge useful discussions with A. Kahn and R. L. Gunshor. Two of the authors (S.H. and J.J.P.) acknowledge support from the Commission of the European Union under the Human Capital and Mobility Program.

¹See, for example, Proc. SPIE **2346**, 1 (1994), and references therein.

²S. Guha, J. M. DePuydt, J. Qiu, G. E. Höfler, M. A. Haase, B. J. Wu, and H. Cheng, Appl. Phys. Lett. **63**, 3023 (1993).

³S. Guha, J. M. DePuydt, M. A. Haase, J. Qiu, and H. Cheng, Appl. Phys. Lett. **63**, 3107 (1993).

⁴J. Petruzzello, B. Greenberg, D. Cammack, and R. Dalby, J. Appl. Phys. **63**, 2299 (1988).

⁵J. L. Batstone, J. W. Steeds, and P. J. Wright, Philos. Mag. **66**, 609 (1992).

⁶L. H. Kuo, L. Salamanca-Riba, J. M. DePuydt, H. Cheng, and J. Qiu, Appl. Phys. Lett. **63**, 3197 (1993); L. H. Kuo, L. Salamanca-Riba, J. M. DePuydt, H. Cheng, and J. Qiu, Philos. Mag. A **69**, 301 (1994).

⁷L. H. Kuo, L. Salamanca-Riba, G. Hofer, B. J. Wu, and M. A. Haase, Proc. SPIE **2228**, 144 (1994).

⁸J. Tanimura, O. Wada, T. Ogama, Y. Endoh, and M. Imaizumi, J. Appl. Phys. **77**, 6223 (1995).

⁹J. M. Gaines, J. Petruzzello, and B. Greenberg, J. Appl. Phys. **73**, 2835 (1993).

¹⁰S. Guha, H. Munkata, and L. L. Chang, J. Appl. Phys. **73**, 2294 (1993).

¹¹E. D. Bourret-Courchesne, Appl. Phys. Lett. **68**, 1675 (1996).

¹²R. Nicolini, L. Vanzetti, Guido Mula, G. Bratina, L. Sorba, A. Franciosi, M. Peressi, S. Baroni, R. Resta, A. Baldereschi, J. E. Angelo, and W. W. Gerberich, Phys. Rev. Lett. **72**, 294 (1994).

¹³A. Bonanni, L. Vanzetti, L. Sorba, A. Franciosi, M. Lomascolo, P. Prete, and R. Cingolani, Appl. Phys. Lett. **66**, 1092 (1995).

¹⁴V. Pellegrini, M. Börger, M. Lazzeri, F. Beltram, J. J. Paggel, L. Sorba, S. Rubini, M. Lazzarino, A. Franciosi, J.-M. Bonard, and J.-D. Ganière, Appl. Phys. Lett. **69**, 3233 (1996).

¹⁵L. H. Kuo, K. Kimura, S. Miwa, T. Yasuda, and T. Yao, Appl. Phys. Lett. **69**, 1408 (1996).

¹⁶The CIL technique has been used previously by us to achieve Zn- and Se-rich interface compositions in ZnSe/GaAs(001) heterostructures and to tune the corresponding band offset (see Refs. 12–14). The present work focuses on the structural as opposed to the electronic effect of the CIL.

¹⁷Minimum contrast is expected when $\mathbf{g} \times \mathbf{b} = 0$ for dislocations and $\mathbf{g} \times \mathbf{R} = 0$ or integer for the stacking faults, where \mathbf{g} identifies the vector of the employed reflection, \mathbf{b} the Burgers vector of the partial dislocation, and \mathbf{R} the displacement vector of the stacking fault; see G. Thomas and M. J. Goringe, *Transmission Electron Microscopy of Materials* (Wiley, New York, 1979). More detail is provided in: J.-M. Bonard, J.-D. Ganière, S. Heun, J. J. Paggel, S. Rubini, L. Sorba, and A. Franciosi, Philos. Mag. Lett. **75**, 219 (1997).

¹⁸S. Heun, J. J. Paggel, L. Sorba, S. Rubini, A. Franciosi, J.-M. Bonard, and J.-D. Ganière, Appl. Phys. Lett. **70**, 237 (1997); J.-M. Bonard, Ph.D. thesis, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 1996.

¹⁹See, for example, K. R. Breen, P. N. Uppal, and J. S. Ahearn, J. Vac. Sci. Technol. B **7**, 758 (1989).

²⁰D. Li and M. Pashley, Phys. Rev. B **49**, 13 643 (1994); S. Takatani, T. Kikawa, and M. Nakazawa, *ibid.* **45**, 8498 (1992).

²¹S. Heun, J. J. Paggel, S. Rubini, and A. Franciosi, J. Vac. Sci. Technol. B **14**, 2980 (1996).

²²A. Raisanen, L. J. Brillson, A. Franciosi, R. Nicolini, L. Vanzetti, and L. Sorba, J. Electron. Mater. **24**, 163 (1995); A. Raisanen, L. J. Brillson, L. Vanzetti, L. Sorba, and A. Franciosi, J. Vac. Sci. Technol. A **13**, 690 (1995); A. D. Raisanen, L. J. Brillson, L. Vanzetti, A. Bonanni, and A. Franciosi, Appl. Phys. Lett. **66**, 3301 (1995).

²³G. Bratina, T. Ozzello, and A. Franciosi, J. Vac. Sci. Technol. B **14**, 2967 (1996); J. Vac. Sci. Technol. A **14**, 3135 (1996); G. Bratina, L. Vanzetti, A. Bonanni, L. Sorba, J. J. Paggel, A. Franciosi, T. Peluso, and L. Tapfer, J. Cryst. Growth **159**, 703 (1996).

²⁴D. Li, J. M. Gonsalves, N. Otsuka, J. Qiu, M. Kobayashi, and R. L. Gunshor, Appl. Phys. Lett. **57**, 449 (1990).

²⁵L. H. Kuo, K. Kimura, T. Yasuda, S. Miwa, C. G. Lin, K. Tanaka, and T. Yao, Appl. Phys. Lett. **68**, 2413 (1996).

Papers from the Fourth International Conference on Nanometer-Scale Science and Technology

**8-12 September 1996
Beijing, People's Republic of China**

**Sponsored by
American Vacuum Society
Beijing Laboratory of Vacuum Physics
Chinese Academy of Sciences
Chinese Vacuum Society
National Natural Science Foundation of China
The State Science and Technology Commission of China**

**Special Editors for the Proceedings:
Dawn Bonnell, Terry A. Michalske,
X. Shen, and Gary McGuire**



**Published for the American Vacuum Society by
the American Institute of Physics, New York, 1997**

International Standard Serial Number: 1071-1023

International Standard Book Number: 1-56396-739-1

Copyright 1997 by the American Vacuum Society

Published by the American Institute of Physics
500 Sunnyside Blvd., Woodbury, NY 11797-2999

Printed in the United States of America

ORGANIZING COMMITTEES

INTERNATIONAL STEERING COMMITTEE

V. Alfeev (Russia)	Y. Kuk (Korea)	L. Samuelson (Sweden)
M. Aono (Japan)	J. Murday (USA)	M. Stedman (UK)
H. Craighead (USA)	S. Pang (China)	E. Teague (USA)
H. Fuchs (Germany)	D. Rugar (USA)	H. Tokumoto (Japan)

INTERNATIONAL PROGRAM COMMITTEE

Shijin Pang, Chair, Beijing Laboratory of Vacuum Physics (China)

Chunli Bai (China)	H. Guntherodt (Switzerland)	Richard Palmer (UK)
Dawn Bonnell (USA)	D. H. Huang (Japan)	Dan Rugar (USA)
Larry Bottomley (USA)	A. Koma (Japan)	T. Sakurai (Japan)
R. Clarke (Australia)	Y. Kuk (Korea)	L. Samuelson (Sweden)
Richard Colton (USA)	C. R. K. Marrian (USA)	H. Tokumoto (Japan)
John Dagata (USA)	M. Martsenyuk (Russia)	K. Urban (Germany)
Liz Dobisz (USA)	Ralph Merkle (USA)	K. Wickramasinghe (USA)
Harold Fuchs (Germany)	Silvio Modesti (Italy)	R. Wiesendanger (Germany)
B. Geerligs (Netherlands)	Lars Montelius (Sweden)	Weisheng Yang (China)
Jean-Claude Gery (France)	J. Murday (USA)	Lide Zhang (China)
H. Gleiter (Germany)	Yuichiro Nishina (Japan)	Zhenyu Zhang (USA)

ORGANIZING COMMITTEE

Pengju Xiang, Co-Chair, Peking Univ. (China)

Zengquan Xue, Co-Chair, CEMS (China)

Marion Churchill (USA)	Dianhong Shen (China)	Sishen Xie (China)
Lili Fang (China)	Xueying Shen (China)	Lei Yuan (China)
Hongfei Liu (China)	Fang-an Shi (China)	Xiufang Zhang (China)
Manu Multani (India)	Jing Wang (China)	Zhijian Zhang (China)

ADVISORY COMMITTEE

Fangyun Chen (China)	Hengde Li (China)	Daheng Wang (China)
Junliang Chen (China)	Lin Li (China)	Shouwu Wang (China)
Duan Feng (China)	Quan Lin (China)	Yu Wei (China)
Zizhao Gan (China)	Yongxiang Lu (China)	Quande Wu (China)
Zhongyi Hua (China)	Liqin Shao (China)	Cunhao Zhang (China)
Guofan Jin (China)	Changxu Shi (China)	Guangzhao Zhou (China)
Kexin Kuo (China)	Youqi Tang (China)	

Papers from the Fourth International Conference on Nanometer-Scale Science and Technology

Preface.....	1294
<i>Surface Atomic Structure</i>	
Growth and characterization of Si and Ge clusters on ordered C ₆₀ overlayers D. Klyachko and D. M. Chen.....	1295
Interaction of C ₆₀ with the (3×3) and ($\sqrt{3}\times\sqrt{3}$) surfaces of 6H-SiC(0001): Adsorption, decomposition, and SiC growth L. Li, Y. Hasegawa, H. Shinohara, and T. Sakurai.....	1300
Novel C ₆₀ -DDME complex thin film with electrical bistable properties M. Ouyang, Z. Q. Xue, K. Z. Wang, Q. D. Wu, and D. Qiang.....	1304
Si- and C-rich structure of the 6H-SiC(0001) surface L. Li, Y. Hasegawa, and T. Sakurai.....	1307
Adsorption geometries of 1,9-decadiene on Si(111) 7×7 studied by scanning tunneling microscopy D. Shachal and Y. Manassen.....	1310
Study on surface and interface structures of nanocrystalline silicon by scanning tunneling microscopy Juning Gao, Haiqiang Yang, Ning Liu, Dongxia Shi, Yueshan Jiang, Zengquan Xue, Shijin Pang, and Yuliang He.....	1313
Diffusional attractions between voids on a Si(111)7×7 surface E. Ter Ovanesyan, Y. Manassen, N. Ramesh Rao, and Z. Olami.....	1317
Liquid-phase adsorption process of β -picoline on stilbite (010) observed by atomic force microscopy Masaharu Komiyama and Minming Gu.....	1325
<i>Nanoscale Magnetic Properties</i>	
Magnetic nanostructures studied by scanning probe microscopy and spectroscopy R. Wiesendanger, M. Bode, M. Kleiber, M. Löhndorf, R. Pascal, A. Wadas, and D. Weiss.....	1330
Magnetotransport through mesoscopic antidot arrays Hongqi Xu.....	1335
Magnetic property of NiO ultrafine particles with a small Ni core Sanshiro Sako, Kazunari Ohshima, Masahiro Sakai, and Shunji Bandow.....	1338
Magnetic force microscope images of magnetic domains in magnetic garnet F. Tian, C. Wang, G. Y. Shang, N. X. Wang, and C. L. Bai.....	1343
Field induced antiferromagnetic ordering in NiS cluster confined in zeolite Y Weiping Ding, Wei Zhong, Ning Zhang, Zaibing Guo, Gang Gu, and Youwei Du.....	1347
Study of the complex of the Schiff-base polymer with sulfate iron by magnetic force microscopy F. Tian, C. Wang, C. L. Bai, W. G. Li, and M. X. Wan.....	1350
Imaging, polymerization, and reconstruction of polystyrene films with a scanning tunneling microscope Z. Y. Hua and W. Xu.....	1353
Application of the scanning force microscope in structuring and in temperature-dependent analysis of Au nanostructures H. Göbel, L. Jacobs, and P. von Blanckenhagen.....	1359
<i>Surface Modification and Nanolithography</i>	
Scanning tunneling microscopy modification of Ag thin films on Si(100): Local rearrangement of the Si substrate by Ag/Si eutectic phase formation U. Memmert, U. Hodel, and U. Hartmann.....	1364
Novel lithography and signal processing with water vapor ions H. W. P. Koops, E. Dobisz, and J. Urban.....	1369

(Continued)

Fabrication of surface nanostructures by scanning tunneling microscope induced decomposition of SiH_4 and SiH_2Cl_2	
H. Rauscher, F. Behrendt, and R. J. Behm	1373
Studies of field related effects in the fabrication process on graphite using a scanning tunneling microscope	
Chen Wang, Xiaodong Li, Guangyi Shang, Xiaohui Qiu, and Chunli Bai	1378
Nanostructuring by reactive accelerated cluster erosion	
A. Gruber, J. Gspann, and P. von Blanckenhagen	1382
Molecular patterns by manipulating DNA molecules	
Zhen-qian Ouyang, Jun Hu, Sheng-fu Chen, Jie-lin Sun, and Min-qian Li	1385
Observation of modification and recovery of local properties of polyethylene oxide	
H.-Y. Nie, M. Motomatsu, W. Mizutani, and H. Tokumoto	1388
Scanning probe nanofabrication of chemically active areas on substrate covered with organosilane monolayers	
Hiroyuki Sugimura and Nobuyuki Nakagiri	1394
Fabrication of hybrid superconductor–semiconductor nanostructures by integrated ultraviolet-atomic force microscope lithography	
Pasqualantonio Pingue, Marco Lazzarino, Fabio Beltram, Ciro Cecconi, Paolo Baschieri, Carlo Frediani, and Cesare Ascoli	1398
<i>Single Electron Tunneling</i>	
Metal-based single electron transistors	
W. Chen and H. Ahmed	1402
Electron beam dot lithography for nanometer-scale tunnel junctions using a double-layered inorganic resist	
Satoshi Haraichi, Toshimi Wada, Sucheta M. Gorwadkar, Kenichi Ishii, and Hiroshi Hiroshima	1406
Negative differential resistance on single electron transport in a junction array of ultrasmall islands	
Hiroshi Nakashima and Kiyohiko Uozumi	1411
<i>Self-Assembly and Self-Organization</i>	
Patterning of Langmuir–Blodgett film with ultrahigh vacuum-scanning tunneling microscope/atomic force microscope	
H. Hamanaka, T. Ono, and M. Esashi	1414
Molecular organization of azobenzene derivatives at the liquid/graphite interface observed with scanning tunneling microscopy	
P. C. M. Grim, P. Vanoppen, M. Rücker, S. De Feyter, S. Valiyaveetil, G. Moessner, K. Müllen, and F. C. De Schryver	1419
Formation of nanocolumn self-assembly by solvent polarity control	
Xiangdong Chai, Wensheng Yang, Yunwei Cao, Yueshun Jiang, Ran Lu, Tiejin Li, Huixin He, Yongqiang Wang, and Zhongfan Liu	1425
Writing and reading bit arrays for information storage using conductance change of a Langmuir–Blodgett film induced by scanning tunneling microscopy	
K. Takimoto, R. Kuroda, S. Shido, S. Yasuda, H. Matsuda, K. Eguchi, and T. Nakagiri	1429
Semiconducting, gas-sensing properties of Europium bisphthalocyanine Langmuir–Blodgett thin films	
Bingjie Liang, Chunwei Yuan, and Yu Wei	1432
<i>Nanowires and Nanoparticles</i>	
Formation of Si nanowire by atomic manipulation with a high temperature scanning tunneling microscope	
R. Hasunuma, T. Komeda, H. Mukaida, and H. Tokumoto	1437
Study on photoelectric properties of a TiO_2 nanoparticle	
Yongmei Chen, Yaan Cao, Yubai Bai, Wensheng Yang, Jihua Yang, Haiyan Jin, and Tiejin Li	1442
Formation of nanocrystals in <i>a</i>-Si thin films induced by pulsed laser ultraviolet irradiation	
Z. Chvoj, V. Cháb, and O. Borusík	1445
Atomic force microscopy investigations on the surface topographies of aluminum-based composite containing nanocluster diamond	
Q. Ouyang, B. Wang, and K. Okada	1449

(Continued)

Calculation of electronic energy levels in artificially confined cavities of a sphere and a circular cylinder Yong Fang Zhao, Xiao Gong Jing, Lin Song Li, Li Jun Wang, Zheng Hui, and Tie Jin Li	1452
Preparation and structural characterization of nanostructured CoAg granular films H. Sang, G. Ni, J. Lu, H. Chen, J. R. Zhang, and Y. W. Du	1456
Fabrication, characterization, and the photoelectric conversion of the nanostructured TiO₂ electrode Huihua Deng, Zuhong Lu, Haifang Mao, and Huijun Xu	1460
Surface vibrational mode of ZnS nanoparticles Jianfeng Xu, Haitao Mao, Yue Sun, and Youwei Du	1465
Photovoltaic study of nanocrystalline TiO₂ film modified with dye molecules Jinghui Fang, Jingwen Wu, Xiaomei Lu, Jianhua Gu, and Zuhong Lu	1468

Optical Responses

Visible luminescence from Si/SiO₂ superlattices S. V. Novikov, J. Sinkkonen, O. Kilpelä, and S. V. Gastev	1471
Numerical study of optical transmission in sandwich slabs: Implication to photon scanning tunneling microscopy Shu Wang, Mufei Xiao, and Jesús M. Siqueiros	1474

Quantitative Topographic Analyses

Analysis of frictional-force image patterns of a graphite surface Naruo Sasaki, Masaru Tsukada, Satoru Fujisawa, Yasuhiro Sugawara, Seizo Morita, and Katsuyoshi Kobayashi	1479
Quantification of topographic structure by scanning probe microscopy James D. Kiely and Dawn A. Bonnell	1483
Error budget of step height and pitch measurement using a scanning tunneling microscope with a three-dimensional interferometer Toru Fujii, Masataka Yamaguchi, Masatoshi Suzuki, Hirofumi Yamada, and Kan Nakayama	1494
Principle of atomic grating and its application in nanotechnology H. W. Hao, H. Qiao, H. K. Cheng, W. W. Wei, and W. K. Yi	1498
Size and arrangement of elementary fibrils in crystalline cellulose studied with scanning tunneling microscopy Y. Z. Zhang, X. L. Chen, J. Liu, P. J. Gao, D. X. Shi, and S. J. Pang	1502
Nanomechanical surface characterization by atomic force acoustic microscopy U. Rabe, V. Scherer, S. Hirsekorn, and W. Arnold	1506

Novel Instrumentation

Detection mechanism of an optical evanescent field using a noncontact mode atomic force microscope with a frequency modulation detection method Masayuki Abe, Takayuki Uchihashi, Masahiro Ohta, Hitoshi Ueyama, Yasuhiro Sugawara, and Seizo Morita	1512
Fabrication of probe tips for reflection scanning near-field optical microscopes: Chemical etching and heating-pulling methods Mufei Xiao, Jesús Nieto, Roberto Machorro, Jesús Siqueiros, and Hector Escamilla	1516
Near-field optical microscopy with uncoated tips: Calibration, chemical contrast on organic crystals, and photolithography G. Kaupp, A. Herrmann, and M. Haak	1521
Preparation of probe tips with well-defined spherical apexes for quantitative scanning force spectroscopy U. D. Schwarz, O. Zwörner, P. Köster, and R. Wiesendanger	1527
Fabrication of a Si scanning probe microscopy tip with an ultrahigh vacuum-scanning tunneling microscope/atomic force microscope Takahito Ono, Hiroaki Saitoh, and Masayoshi Esashi	1531
Conductive supertips for scanning probe applications C. Schössler, J. Urban, and H. W. P. Koops	1535
Simultaneous optical detection techniques, interferometry, and optical beam deflection for dynamic mode control of scanning force microscopy M. Houmady, E. Farnault, T. Yahiro, and H. Kawakatsu	1539

Development of ultrahigh vacuum-atomic force microscopy with frequency modulation detection and its application to electrostatic force measurement	
Takayuki Uchihashi, Masahiro Ohta, Yasuhiro Sugawara, Yoshio Yanase, Tatsuhiko Sigematsu, Mineharu Suzuki, and Seizo Morita	1543
Development of a metal patterned cantilever for scanning capacitance microscopy and its application to the observation of semiconductor devices	
Takuma Yamamoto, Yoshihiko Suzuki, Masayuki Miyashita, Hiroyuki Sugimura, and Nobuyuki Nakagiri	1547
Novel high vacuum scanning force microscope using a piezoelectric cantilever and the phase detection method	
Jiaru Chu, Toshihiro Itoh, Chengkuo Lee, Tadatomo Suga, and Kazutoshi Watanabe	1551
New technique for nanocantilever fabrication based on local electrochemical etching: Applications to scanning force microscopy	
M. Hoummady, E. Farnault, H. Fujita, H. Kawakatsu, and T. Masuzawa	1556
Development of a piezoelectric self-excitation and self-detection mechanism in PZT microcantilevers for dynamic scanning force microscopy in liquid	
Chengkuo Lee, Toshihiro Itoh, Takahiro Ohashi, Ryutaro Maeda, and Tadatomo Suga	1559
Interface imaging by second-harmonic microscopy	
M. Flörsheimer, M. Bösch, Ch. Brillert, M. Wierschem, and H. Fuchs	1564
Scanning acoustic tunneling microscopy and spectroscopy: A probing tool for acoustic surface oscillations	
T. Hesjedal, E. Chilla, and H.-J. Fröhlich	1569
Investigation on a novel vacuum microelectronic pressure sensor with stepped field emission array	
S. H. Xia, J. Liu, D. F. Cui, J. H. Han, S. F. Chen, and L. Wang	1573
Robotic nanomanipulation with a scanning probe microscope in a networked computing environment	
C. Baur, B. C. Gazen, B. Koel, T. R. Ramachandran, A. A. G. Requicha, and L. Zini	1577
Using a new kind of organic complex system of electrical bistability for ultrahigh density data storage	
H. J. Gao, L. P. Ma, H. X. Zhang, H. Y. Chen, Z. Q. Xue, and S. J. Pang	1581
Atomic force microscope-based data storage using replicated media	
B. D. Terris, S. A. Rishton, H. J. Mamin, M. E. Best, J. A. Logan, and D. Rugar	1584
AUTHOR INDEX	1588

PREFACE

The Fourth International Conference on Nanometer Scale Science and Technology (NANO IV) was held in Beijing, China, 8–12 September 1996. The conference was organized under the auspices of the Beijing Laboratory of Vacuum Physics, the Chinese Academy of Sciences, the State Science and Technology Commission of China, the National Natural Science Foundation of China, the American Vacuum Society (AVS), and the Chinese Vacuum Society. Both the Chinese and the American Vacuum Society were sponsoring organizations with the AVS providing meeting management assistance.

The conference was devoted to physical and chemical phenomena which occur on nanometer scales and to their applications in science and technology, and provided superb vignettes of state-of-the-art nanofabrication, scanning probes, and interesting properties of nanostructures.

There were 30 invited talks and 356 papers presented—145 from China and 211 from around the globe, with Japan, Germany, and the U.S. as the principal contributors. This proceedings volume contains papers presented at the NANO-IV, which were reviewed after the meeting.

We thank all the speakers for their presentations, and the reviewers and session chairmen for their efforts. We should also like to commend the authors of the poster presentations for the high standard of this year's contributions.

The generous financial support from both public and private sponsors of the conference and the proceedings is gratefully acknowledged.

We thank Jim Murday and Marion Churchill for their valuable advice and assistance in the conference management. We would also like to express our appreciation of the work of Dawn Bonnell and Terry Michalske, as the conference editors for paper selection and review.

Professor Shijin Pang
Chair, NANO 4 International Program Committee
Director, Beijing Laboratory of Vacuum Physics

Growth and characterization of Si and Ge clusters on ordered C₆₀ overlayers

D. Klyachko^{a)} and D. M. Chen^{b)}

The Rowland Institute for Science, Cambridge, Massachusetts 02142

(Received 12 September 1996; accepted 31 March 1997)

We use *in situ* fabricated ordered C₆₀ overlayers on Si and Ge surfaces as templates to grow isolated Ge and Si clusters, respectively. The growth processes are monitored and characterized by scanning tunneling microscopy, Auger spectroscopy, and low energy electron diffraction. Si and Ge clusters nucleate in the interstices of the C₆₀ layer, and reach a typical lateral size of ~ 10 Å before coalescing. Auger spectra of small Si clusters show an enhanced contributions from states of *s* symmetry, reflecting nontetrahedral coordination of Si atoms in clusters. Room temperature tunneling spectroscopy of isolated clusters reveals an enlarged band gap consistent with the effect of quantum confinement, and a quantized conductance attributable to the Coulomb blockade phenomena in a tip/cluster/substrate double barrier junction. © 1997 American Vacuum Society. [S0734-211X(97)03904-8]

I. INTRODUCTION

The making of self-organized nanoscale structures on substrates with patterned nucleation sites such as miscut crystalline surfaces has been an active endeavor in current nanoscience research. Here we propose a novel concept using highly corrugated surfaces of the crystalline overlayer of large molecules as templates for growing clusters or quantum dots. By virtue of the large lattice constant of molecular films and their high surface corrugation, self-organization of atomic species at the open interstices can be potentially realized. In addition, a semiconducting or a nonconducting molecular template could serve naturally as an electron energy barrier, a condition required for charge confinement in quantum dots.

The remarkable stability of C₆₀ molecules and the high crystallinity and the wide band gap (2.5 eV) of a C₆₀ solid¹ render C₆₀ lattices an ideal choice for testing this concept. Despite being a van der Waals solid, the cohesive energy of a face-centered-cubic (fcc) C₆₀ crystal is ~ 0.28 eV per C₆₀-C₆₀ bond,² and gives rise to considerable crystal strength. C₆₀ has a closed shell electronic configuration that contributes to its molecular stability as well as to the inertness of the C₆₀ crystalline surface. The constituent carbon atoms of C₆₀ are arranged on a spherical shell, hence the C₆₀ crystal surface is highly corrugated. A close match in the bond angle and the bond length between the carbon atoms of C₆₀ and the deposited species in order to form a wetting layer is difficult to achieve. Thus a C₆₀ crystal surface is potentially an ideal substrate for growing zero dimensional nanostructures.

A hypothetical scheme for growing Si clusters on a C₆₀ surface is illustrated in Fig. 1. The first step is to make an ordered C₆₀ overlayer on a Ge(100) substrate, that in turn

serves as a template for the growth of isolated Si clusters. The relatively deep interstices in the C₆₀ layer should play the role of nucleation centers [Fig. 1(a)], just as the steps on an atomic surface do. As more atoms arrive on the surface, isolated Si clusters will start to grow from the seed atoms in the interstices [Fig. 1(b)]. Given the 10 Å separation, a cluster can increase its size to ~ 30 atoms before the nearby clusters coalesce into bigger clusters or islands [Fig. 1(c)].

Below we present our experimental verification of these ideas. Two cluster/substrate compositions were investigated using scanning tunneling microscopy (STM), Auger electron spectroscopy (AES), and low energy electron diffraction (LEED). The first is Si clusters on an ordered monolayer of C₆₀ on Ge(100)2×1, and the second is Ge clusters on C₆₀ crystallites grown on a Si(100) substrate. These sandwichlike arrangements allow us to carefully monitor the coverage, the growth dynamics, and possible contamination, as to be discussed in detail later in the article.

II. GROWTH PROCEDURES

The experiments are carried out in an ultrahigh vacuum (UHV) system with a base pressure of 5×10^{-11} Torr. For a typical specimen fabrication, we first prepare clean Si and Ge surfaces by ion sputtering followed by annealing to 1100 and 760 °C, respectively. C₆₀ is then deposited from a Knudsen cell held at 500 °C onto the clean Si substrate held at 100–150 °C, or onto the Ge at 80–100 °C. The long range order of the C₆₀ overlayer on Ge(100) or the crystallites on Si(100) are then verified by LEED and STM. Once the C₆₀ deposition reaches the desired coverage, we switch to deposition of the semiconductor species. The Ge is deposited from a molecular beam epitaxy (MBE) cell heated to 960 °C, whereas the silicon beam is produced from direct current heating of a thin Si strip positioned near the substrate. The typical deposition rate of the Si or Ge is ~ 0.02 monolayers (ML)/min. The substrates are not heated during cluster growth, but a slight rise in the surface temperature might occur due to thermal radiation heating from the Si strip.

^{a)}Current address: Dept. of Nuclear Medicine, University of Sherbrooke, Quebec J1H5N4, Canada.

^{b)}Author to whom correspondence should be addressed; Electronic mail: chen@risvax.rowland.org

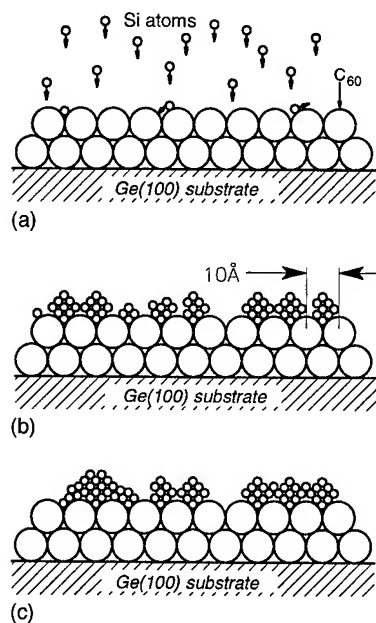


FIG. 1. A schematic showing the concept of using ordered C_{60} overlayers as a template to grow isolated atomic clusters. (a) Nucleation at interstices; (b) forming isolated clusters; (c) coalescence.

The growth and detailed structure of a monolayer C_{60} on $Ge(100)2 \times 1$ and of the C_{60} crystallites on Si have been studied well and are reported elsewhere.³⁻⁵ In the remainder of this article, we shall focus on the growth and characterization of the clusters.

III. AUGER CHARACTERIZATION

In the previously discussed two step deposition processes, AES is used to monitor the coverage and to detect possible impurities. All the cluster samples used for subsequent STM studies were checked to ensure that they were free of contamination. Figure 2(a) is a typical differential AES spectrum of a relatively high coverage of Si on a $C_{60}/Ge(100)$ substrate. It revealed only the signals from Si, C, and Ge, indicating that a clean deposition process was achieved.

During growth of Si clusters, we also recorded changes in the shape of the Si LVV Auger line as the cluster sizes increase with the deposition time. As an example, Fig. 2(b) shows the evolution of the Si LVV line shape, along with a reference spectrum taken on a clean Si(100) surface. The cluster spectra are obtained by subtracting a spectrum taken on a $C_{60}/Ge(100)$ substrate from the spectra for Si on $C_{60}/Ge(100)$. The resultant spectra were normalized to the amplitude of the Si LVV peak. The inelastic background under the peak was interpolated and removed. The shape of the Si LVV spectrum bears the features of the self-convolution of the density of states in the Si valence band.⁶ The latter includes lower energy s states and higher energy p states. The three Auger peaks at 70, 82, and 89 eV in the Si LVV spectra [Fig. 2(b)] are associated with the LV_sV_s , LV_sV_p , and LV_pV_p transitions, respectively.⁶ In Fig. 2(b) the LV_sV_s and LV_sV_p contributions are most pronounced at lower coverage (i.e., for small clusters), both decrease slowly

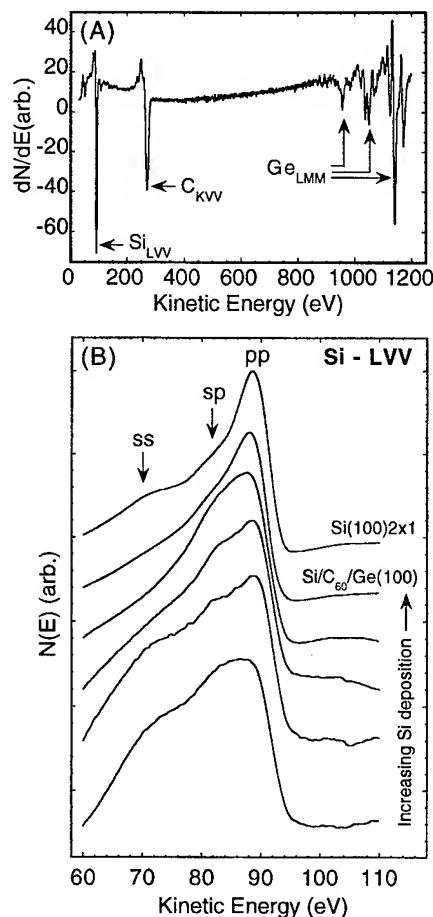


FIG. 2. Auger spectra of Si clusters on C_{60} overlayers that are grown on the $Ge(100)2 \times 1$ substrate. The specimen is prepared by a two step molecular beam deposition process.

as Si deposition progresses, and eventually they approach the level of amorphous Si. The decrease of LV_sV_s and LV_sV_p contributions to the Si LVV spectrum from a cluster to a bulk film can be attributed to the increase of the sp hybridization associated with the transition to tetrahedral bonding in the bulk Si.⁷ The s wave is located closer to the atomic core than the p wave and this leads to an enhancement of the LV_sV_s and LV_sV_p transitions in the clusters.

IV. SCANNING TUNNELING MICROSCOPY OF Si/Ge CLUSTERS

A. Si clusters on $C_{60}/Ge(100)$

Figure 3 shows typical STM images of a monolayer C_{60} on $Ge(100)$ before and after Si deposition. The original C_{60} lattice [Fig. 3(a)] is essentially free of defects except for a few C_{70} impurities (the brighter molecule). After the surface is exposed briefly to the Si beam, the STM images [Figs. 3(b) and 3(c)] show many isolated protrusions with various sizes. Since AES spectra confirm that there is no contamination in our sample and that the C_{60} layer clearly remains intact, we conclude that *all* the bright protrusions are either Si atoms or Si clusters. Figure 3(c) shows that the diffusion

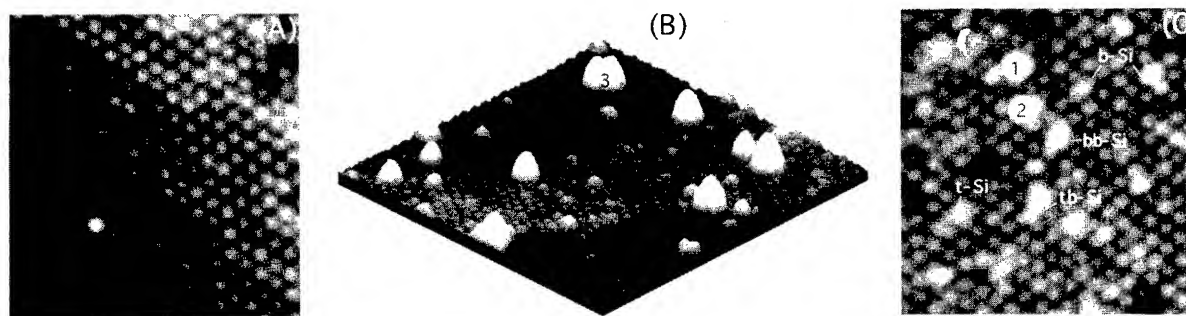


FIG. 3. (a) A $160 \text{ Å} \times 160 \text{ Å}$ STM image of a rhombic C_{60} lattice on the $\text{Ge}(100)2 \times 1$; (b) and (c) a $200 \text{ Å} \times 200 \text{ Å}$ 3D view and a $160 \text{ Å} \times 160 \text{ Å}$ plane view, respectively, of the STM images taken after a monolayer C_{60} on $\text{Ge}(100)2 \times 1$ was exposed briefly to an atomic beam of Si. STM images were taken with a tip bias of -3 V and a current set at 40 pA .

length of Si atoms on the C_{60} layers is quite small, and that nucleation takes place at either the bridge site (b-Si) or the triangular site (t-Si). We suppose that Si occupation in two adjacent bridge sites (bb-Si) or in a bridge and a triangular site is the natural pathway to the formation of the Si clusters (marked as 1 and 2). Note that cluster 3 in Fig. 3(b) appears to be the beginning of a merger of two isolated clusters separated by only $\sim 10 \text{ Å}$.

B. Ge clusters on $C_{60}/\text{Si}(100)$

Ge clusters are grown on the (111) plane of C_{60} crystallites which in turn are supported by a $\text{Si}(100)$ substrate. Figure 4(a) shows a three-dimensional (3D) perspective view of a clean C_{60} crystallite. The top (111) facet, or growth plane, is a near perfect close-packed hexagonal lattice similar to that in Fig. 3(a). A 3D view of a crystallite after the deposition of Ge is shown in Fig. 4(b) and part of its top view is shown in Fig. 4(c). Although the Ge coverage is much higher than in the preceding Si deposition, the bare region of the top C_{60} lattice and the side facets of the C_{60} crystallite remain unaffected. Judging by the absence of defects in the immediate surroundings of the clusters, we speculate that the C_{60} lattice underneath the Ge clusters also remains intact. It should be noted that, because of the severe convolution effect between the tip and clusters, STM images do not offer true dimensions or a detailed structure of the clusters.

The most significant conclusion drawn from our STM results is that, in spite of the van der Waals intermolecular bonding, the C_{60} lattice and the crystallite have sufficient stability to support the growth of Si and Ge clusters.

V. TUNNELING SPECTROSCOPY OF INDIVIDUAL Ge CLUSTERS

The size distribution is a common problem encountered in the characterization of the physical properties of nanoscale structures using nonlocal probing techniques such as optical fluorescence measurement. With a STM, this problem can be easily overcome because of the ability to select individual clusters in a well controlled experiment such as the one described above. We have performed extensive tunneling spectroscopy ($I-V$ measurements) on numerous isolated clusters.

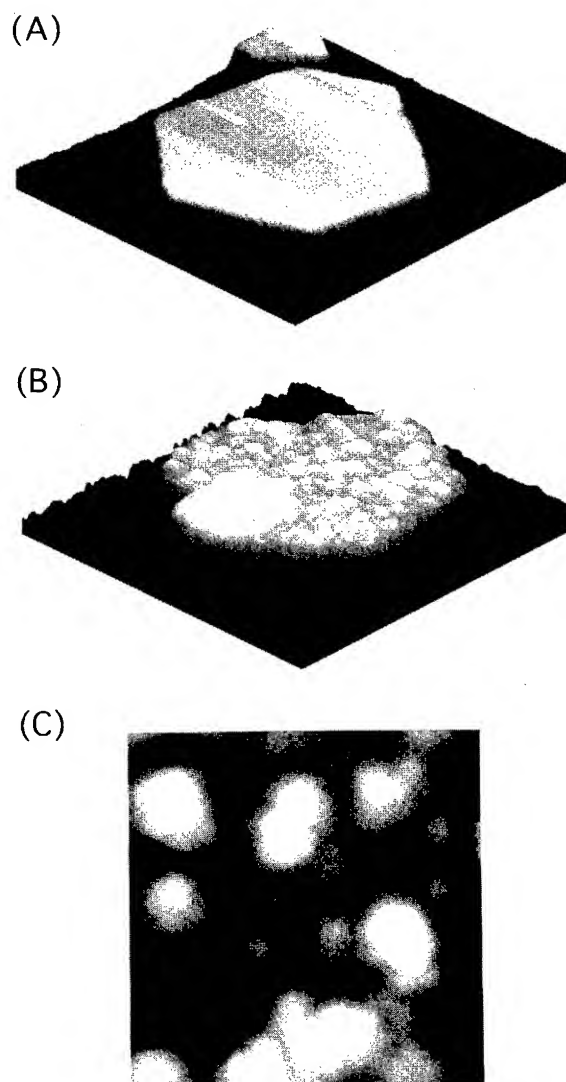


FIG. 4. (a) A 3D rendition of a STM image showing a C_{60} crystallite 156 Å high (18 layers) and $\sim 1800 \text{ Å}$ wide. The top surface is a close-packed hexagonal lattice with a lattice constant of $\sim 10 \text{ Å}$, i.e., a (111) plane of a fcc C_{60} crystal; (b) a STM image of a similar C_{60} crystallite after Ge deposition. Its height is 147 Å (17 layers) and its width is $\sim 800 \text{ Å}$. (c) A $150 \text{ Å} \times 150 \text{ Å}$ STM image taken on top of crystallite (b). The images were acquired with a tip bias of -5 V and a current set at 20 pA .

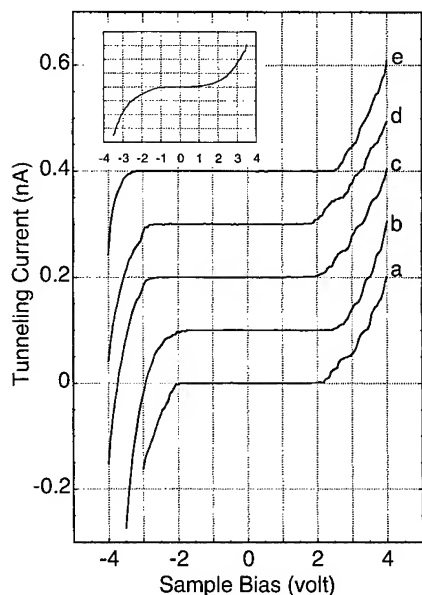


FIG. 5. I - V spectra of five isolated Ge clusters grown on the (111) surface of a C_{60} crystallite that is supported by a Si(100) substrate. The inset is a spectrum taken on a clean C_{60} overlayer.

Figure 5 displays the I - V spectra taken on five different Ge clusters with sizes ranging from 30 to 50 Å. Two characteristic features are quite visible on all the curves: (i) large band gaps [spectrum (a)] and (ii) steps in the tunnel current [spectrum (b)].

The large band gap, 4–5.5 eV, is indicative of the quantum confinement in nanoscale structures. The actual size of a cluster is less than 50 Å, so the measured gap is comparable to the mean field theory calculation.⁸ Nonetheless, two other effects need to be considered as well. First, the apex of the STM tip could adsorb a C_{60} molecule or a Ge cluster. If this is the case, the true gap should be roughly one-half that measured in the I - V spectrum. Although a great deal of care was exercised to avoid this situation while acquiring these I - V spectra, we cannot exclude this possibility absolutely. Second, a voltage drop across the Si- C_{60} interface or the C_{60} crystallite could also lead to inflated values for the gap (or the width of a Coulomb staircase; see below). Clearly, much more work is needed to elucidate the nature of the large values for the band gap measured.

To address the stepped conductance at higher bias, in Fig. 6 we expand spectrum (b) of Fig. 5 in the region of the 2–4 V bias, and display the derivative of the whole (b) spectrum in the inset. The steps are nearly equidistant in both width (~ 300 mV) and height (~ 47 pA), reminiscent of the Coulomb staircase observed in single electron tunneling junctions.⁹ The step width and height of a Coulomb staircase is equal to e/C and e/RC , respectively, where C is the dominating junction capacitance and R the resistance. Using the value from spectrum (b) we obtain $C \approx 5 \times 10^{-19}$ F, and $R \approx 6 \times 10^9 \Omega$. For a spherical conductor of 40 Å diameter placed between two parallel plates 20 Å apart, $C \approx 3 \times 10^{-19}$ F. From this comparison, it is tempting to attribute

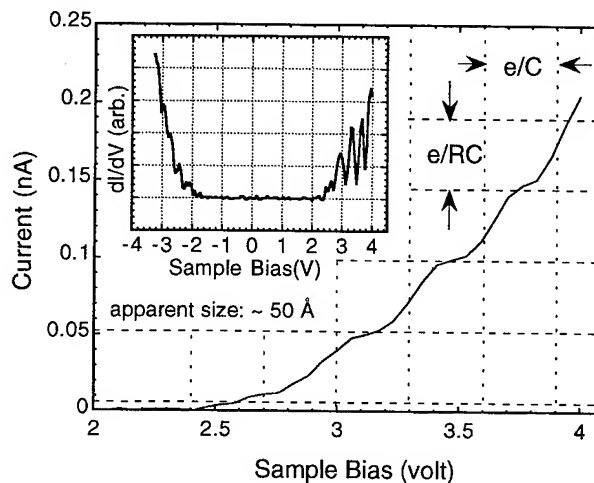


FIG. 6. An expanded view from the 2–4 V bias range of I - V spectrum (b) shown in Fig. 5. The inset displays the derivative of the whole (b) spectrum of Fig. 5.

the observed steplike I - V spectrum to the phenomenon of Coulomb blockade. Room temperature single electron tunneling junction involving a STM probe has been reported previously.^{10,11} A more adequate proof for such a conclusion, however, requires a series of I - V measurements on clusters (quantum dots) of variable sizes. We do not have sufficient statistics at the present time. Much more effort is needed in this direction.

VI. CONCLUSIONS

We have proposed a novel concept of using lattices of large molecules as a template for growing atomic clusters (or quantum dots). Our experiment shows that isolated Si and Ge clusters from a few atoms up to ~ 50 Å in size can be grown successfully using this method. Tunneling spectroscopy performed on individual clusters reveals quantized conductance and an enlarged energy band gap, which are interpreted by the effects of the Coulomb blockade and quantum confinement, respectively. Our preliminary experiments show that metallic dots can also be fabricated using this technique.¹² It is obvious that this concept could be extended to include other types of molecules. Of particular interest are those that provide greater stability and hence permit higher cluster growth temperature. Despite our initial success in implementing this new concept, its practical value ultimately depends on the possibility of preparing clusters (quantum dots) that can be distributed uniformly and orderly on these templates, a challenge seemingly formidable, yet worth taking on.

ACKNOWLEDGMENTS

The authors thank L. Ruan for providing supporting evidence for our conclusions and A. Bass for helpful discussions. This work was supported by the Rowland Institute for Science.

- ¹J. H. Weaver and D. M. Poirier, in *Solid State Physics*, edited by H. Ehrenreich and F. Spaepen (Academic, New York, 1994), Vol. 48, p. 1.
- ²A. V. Hamza and M. Balooch, *Chem. Phys. Lett.* **186**, 603 (1992).
- ³D. Klyachko and D. M. Chen, *Phys. Rev. Lett.* **75**, 3693 (1995).
- ⁴D. M. Chen, in *Recent Advances in the Chemistry and Physics of Fullerenes and Related Materials*, edited by R. S. Ruoff and K. M. Kadish (The Electrochemical Society, Reno, NV, 1995), p. 1311.
- ⁵X. D. Wang *et al.*, *Phys. Rev. B* **47**, 15 923 (1993).
- ⁶S. M. Durbin and T. Gog, *Phys. Rev. Lett.* **63**, 1304 (1989).
- ⁷W. L. Brown *et al.*, *Science* **235**, 860 (1987).
- ⁸T. Takagahara and K. Takeda, *Phys. Rev. B* **46**, 15 578 (1992).
- ⁹*NATO Advanced Study Institute on Single Charge Tunneling*, edited by H. Grabert and M. H. Devoret (Plenum, New York, 1991).
- ¹⁰P. Radojkovic *et al.*, *Surf. Sci.* **361–362**, 890 (1996).
- ¹¹R. P. Andres *et al.*, *Science* **272**, 1323 (1996).
- ¹²L. Ruan and D. M. Chen (unpublished data, 1996).

Interaction of C_{60} with the (3×3) and $(\sqrt{3}\times\sqrt{3})$ surfaces of 6H-SiC(0001): Adsorption, decomposition, and SiC growth

L. Li^{a)} and Y. Hasegawa

Institute for Materials Research, Tohoku University, Sendai 980-77, Japan

H. Shinohara

Chemistry Department, Nagoya University, Nagoya 464-01, Japan

T. Sakurai

Institute for Materials Research, Tohoku University, Sendai 980-77, Japan

(Received 8 September 1996; accepted 18 March 1997)

Scanning tunneling microscopy was used to study the interaction of C_{60} with the silicon rich (3×3) and carbon rich $(\sqrt{3}\times\sqrt{3})$ surfaces of 6H-SiC(0001). For both reconstructions, triangular or hexagonal islands commensurate with the substrate structure were observed at submonolayer coverages. The C_{60} exhibits island growth on these surfaces, with the second layer nucleating on the C_{60} islands before completion of the first layer. For coverages greater than one monolayer, a closed-packed face-centered-cubic (111) structure was formed on the (3×3) surface. On the $(\sqrt{3}\times\sqrt{3})$ surface, a disordered layer was observed. A SiC film was also grown using C_{60} as the carbon source. © 1997 American Vacuum Society. [S0734-211X(97)05404-8]

I. INTRODUCTION

Since it was first synthesized in 1984, the unique properties of the C_{60} molecule has attracted great attention.^{1,2} Fullerene research has emphasized the building of increasingly elaborate derivatives of the archetypal C_{60} molecule. A significant effort, however, has also been devoted to studies of the adsorption and film growth of the C_{60} molecule on various semiconductor surfaces using many techniques, including scanning tunneling microscopy (STM).³ An important aspect of the research is an understanding of the adsorption geometry in the submonolayer regime and mechanisms of thin film growth in the multilayer regime, as well as the interaction between the C_{60} molecule and the substrate. In most of the cases studied, the structure of the monolayer and multilayer films consists of a hexagonal, close-packed (111) layer. One interesting exception that was reported is the formation of a highly strained (110) oriented C_{60} film on the GaAs(001) As-rich 2×4 surface.⁴ Several groups were interested in the potential industrial applications of fullerene,^{5,6} particularly with regard to the decomposition of C_{60} on Si surfaces in order to grow SiC films. Ion-beam deposition of hydrogen-free, hard amorphous carbon films from a fullerene discharge was also reported.⁷

In this work, we present a STM study of the interaction of C_{60} with 6H-SiC(0001) (3×3) and $(\sqrt{3}\times\sqrt{3})$ surfaces. The 6H-SiC(0001) surface exhibits various phases, such as $(\sqrt{3}\times\sqrt{3})$, (2×3) , $(2\sqrt{3}\times6\sqrt{3})$, and (3×3) .^{8,9} Among them, the $(\sqrt{3}\times\sqrt{3})$ structure is carbon rich and has very low corrugation (0.2 Å), while the (3×3) structure is Si rich and has large corrugation (2.5 Å). Interesting adsorption behavior can be expected for these two surfaces. This is similar to the case of the GaAs(001) surface where C_{60} films grown on

(2×6) and (2×4) surfaces showed different structures.^{4,10} Furthermore, C_{60} is easily sublimated at around 550 °C, and has a vapor pressure of 2×10^{-3} Torr.¹¹ Thus C_{60} could be used as a clean and convenient source of carbon for the growth of SiC. It has several advantages over carbon produced by decomposing hydrocarbon gases and vaporizing graphite. For example, it eliminates the incorporation of hydrogen and hydrocarbon fragments into the film which occurs in gas source molecular beam epitaxy (MBE). It does not require an elaborate apparatus to vaporize as is necessary for solid source MBE. It also avoids chemical contamination during heating of the carbon source due to its low sublimation temperature. Therefore, it is of great technological interest to study the interaction of C_{60} with SiC surfaces and to understand the decomposition mechanism on SiC surfaces at elevated temperatures.

II. EXPERIMENT

Details of the STM instrumentation and cleaning procedure of the SiC sample have been reported elsewhere.^{9,12} Briefly, our "FI-STM" is a combination of STM and a room temperature field ion microscope (FIM) which was used for precise characterization of the scanning tip. The (99.95% purity) C_{60} powder was put into a small Ta dispenser. Sublimation of the C_{60} molecules was performed by directly heating the C_{60} doser at 340 °C. The *n*-type 6H-SiC(0001) single crystals were supplied by Cree Research. Clean SiC surfaces were obtained by annealing the sample at 900–1000 °C in a flux of Si, which was produced by resistively heating a small Si wafer. The STM imaging was performed at room temperature using the constant current mode with sample bias between –1.2 and –2.5 V, unless specified otherwise.

^{a)}Present address: Chemical Engineering Dept., University of California, Los Angeles, CA 90095; Electronic mail: lian@seas.ucla.edu

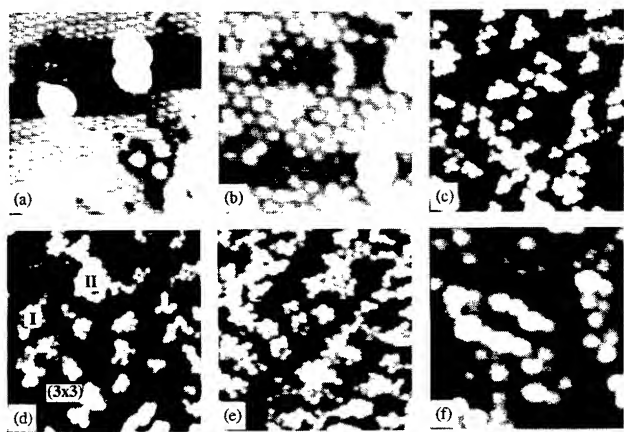


FIG. 1. STM images of the 6H-SiC(0001) surface after deposition of various amounts of C₆₀ at room temperature. (a) The ($\sqrt{3}\times\sqrt{3}$) surface with 0.02 ML coverage. The image size is 80 Å×80 Å, and sample bias is +1.2 V. (b) The (3×3) surface (60 Å×60 Å) with 0.02 ML coverage. (c) The ($\sqrt{3}\times\sqrt{3}$) surface (300 Å×300 Å) with 0.4 ML coverage. (d) The (3×3) surface (320 Å×320 Å) with 0.4 ML coverage. (e) The ($\sqrt{3}\times\sqrt{3}$) surface (320 Å×320 Å) with 1 ML coverage. (f) The internal structure of the C₆₀ on the ($\sqrt{3}\times\sqrt{3}$) surface imaged by STM. The image size is 160 Å×160 Å.

III. RESULTS AND DISCUSSION

Before presenting the results of C₆₀ adsorption on the ($\sqrt{3}\times\sqrt{3}$) and (3×3) surfaces, we wish to briefly mention the atomic structure of the two surfaces. Details of the structure of these surfaces can be found elsewhere.^{8,9} The ($\sqrt{3}\times\sqrt{3}$) surface showed a honeycomb structure in the filled state and a hexagonal structure in the empty state. This surface thus exhibited height contrast reversal in the STM images. The corrugation of the surface is 0.2 Å. In the case of the (3×3) surface, a hexagonal array of protrusions in both the filled and empty states with a corrugation of 2.5 Å forms a (3×3) reconstruction. We now present the results of C₆₀ adsorption on these two surfaces. The ($\sqrt{3}\times\sqrt{3}$) and (3×3) surfaces following an approximately 0.02 monolayer (ML) deposition of C₆₀ are shown in Figs. 1(a) and 1(b), respectively. In Fig. 1(a), the largest bright quasispherical protrusion corresponds to individual C₆₀ molecules. There is a total of three visible, all located in the upper half of the image. Also note that there are several bright and dark spots in the image which are caused by defect sites commonly seen on these two surfaces.^{8,9} The height of the protrusions varies between 6 and 8 Å, depending on the bias voltage. For the (3×3) surface shown in Fig. 1(b), the C₆₀ is adsorbed on top of the protrusions. Again, the C₆₀ molecules are imaged as the largest bright spots, with a total of two being visible in this image. On both surfaces, large area scans indicated that there is no preferential adsorption of C₆₀ molecules at the step edges or at defect sites. While most of the molecules are isolated on the surface, large area images showed that approximately 5% of the C₆₀ molecules formed dimers [see Fig. 1(a)]. However, after the molecules found the most favorable adsorption site they were very stable and showed no sign of diffusion. Even after annealing the surfaces to 300–400 °C, the population of the C₆₀ dimers did not change.

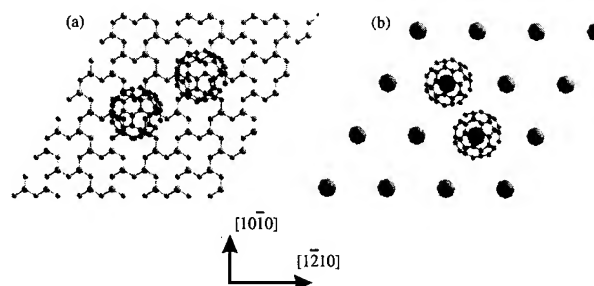


FIG. 2. Schematic diagram of the adsorption model for C₆₀ on ($\sqrt{3}\times\sqrt{3}$) and (3×3) surfaces. (a) The C₆₀ adsorbs at the center of the vacancy on the ($\sqrt{3}\times\sqrt{3}$) surface, and forms a dimer with the second nearest neighbor C₆₀. The large and small balls are Si and C atoms, respectively. (b) The C₆₀ adsorbs at the top of the (3×3) spots, and forms a dimer with its nearest neighbor C₆₀.

When the C₆₀ coverage was increased, the molecules nucleated into clusters of either triangular or hexagonal shape. This is seen in Figs. 1(c) and 1(d), where about 0.4 ML of C₆₀ has been deposited. Using a line-scan profile, it can easily be seen that the C₆₀ clusters are commensurate with the substrate structure on both surfaces. In Fig. 1(d), the clean (3×3) domain, the first layer C₆₀, and the second layer C₆₀ are labeled (3×3), I, and II, respectively. One interesting observation at this coverage is that the second layer C₆₀ molecules grow on top of the first layer before completion of the first layer.

The ($\sqrt{3}\times\sqrt{3}$) surface after 1 ML of C₆₀ deposition is shown in Fig. 1(e). Again islandlike growth was observed, producing multilayer islands while at the same time leaving parts of the clean surface exposed. Similar behavior was observed on the (3×3) surface. Another interesting observation was that the C₆₀ intramolecular structure could be imaged for C₆₀ molecules in both the first and second layers. In Fig. 1(f), three parallel stripes are visible within the molecules. These indicate that the rotation of the C₆₀ molecules was arrested by the surface. This intramolecular structure is very similar to that observed for C₆₀ adsorption on Si surfaces in that three or four stripes were also seen in the first layer C₆₀ molecules.

To elucidate the adsorption behavior of the C₆₀ molecule on the 6H-SiC(0001) ($\sqrt{3}\times\sqrt{3}$) and (3×3) surfaces, we show in Fig. 2 a model of the C₆₀ molecules' adsorption site and dimer formation on these two surfaces derived from the STM results. The C₆₀ adsorbs at the center of the honeycomb (protrusion) of the ($\sqrt{3}\times\sqrt{3}$) structure. But a satisfactory atomic model is not yet available for the (3×3) surface; thus only a schematic diagram of the (3×3) geometry is shown. Based on the adsorption model proposed, the formation of C₆₀ dimers on these two surfaces can be understood as follows. When an arriving C₆₀ molecule approaches an existing C₆₀ molecule on the ($\sqrt{3}\times\sqrt{3}$) surface, the best bonding site for the molecule is the second nearest neighbor because of its large diameter (7.1 Å). Thus a C₆₀ dimer will have a separation in the $2\sqrt{3}$ lattice spacing of SiC(0001)-(1×1), i.e., of 10.7 Å. Similar to this, the ideal C₆₀ dimer separation is 9.3 Å on the (3×3) surface. A previous theoretical study has

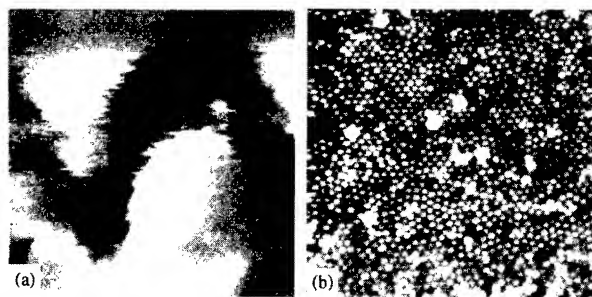


FIG. 3. (a) STM image of the multilayer fcc (111) C₆₀ film grown on the (3×3) surface (160 Å×160 Å). (b) STM image of the (3×3) surface (320 Å×320 Å) after annealing the fcc (111) C₆₀ film at 400 °C, showing a structure with short range ordering.

shown that the equilibrium bond length for a C₆₀ dimer is 9.84 Å.¹³ Compared with the dimer separation in the proposed model, the intermolecular spacing is then about 9.1% larger on the ($\sqrt{3}\times\sqrt{3}$) surface, but about 5.5% smaller on the (3×3) surface. Therefore, C₆₀ dimers are under tensile stress on the ($\sqrt{3}\times\sqrt{3}$) surface whereas they are under compressive stress on the (3×3) surface.

It is of interest to compare our observation of the adsorption behavior of C₆₀ on the ($\sqrt{3}\times\sqrt{3}$) and (3×3) surfaces of 6H-SiC(0001) with the previous reported results on Si(100)-(2×1), Ge(100)-(2×1), and GaAs(001)-(2×4) As surfaces.^{4,14} Substrate structures were found to play a critical role in determining C₆₀ adsorption and film growth. The (2×1) reconstruction of the Si(100) and Ge(100) surfaces forms a trough in between the dimer rows, which confines the C₆₀ molecules so that they are unable to adopt closed hexagonal packing. Instead, the C₆₀ molecules form a *c*(4×3) structure at low coverages.¹⁴ Adsorption of C₆₀ on the anisotropic and strongly corrugated GaAs(001)-(2×4) As surface results in a molecular chain consisting of C₆₀ dimers. This is also confined to the troughs of the substrate at low C₆₀ coverage. This pair-chain structure then serves as a template for the growth of a strained van der Waals C₆₀ face-centered-cubic (fcc) (110) overlayer.⁴ In the present case of ($\sqrt{3}\times\sqrt{3}$) and (3×3) surfaces of 6H-SiC(0001), it also demonstrates that the substrate structures and site-specific C₆₀-substrate interaction play a leading role in adsorption and film growth. The constraints imposed by the substrate result in the formation of a stressed C₆₀ dimer on both surfaces. On the ($\sqrt{3}\times\sqrt{3}$) surface, the C₆₀ dimer lattice contracts with respect to the ideal bulk fcc structure of C₆₀, whereas it expands on the (3×3) surface. Thus, when more C₆₀ molecules are deposited, the resulting triangular or hexagonal C₆₀ clusters are under even larger stress, so that a closed-packed structure cannot be formed at the monolayer coverage regime.

Film growth in the multilayer regime, however, showed different behavior on the ($\sqrt{3}\times\sqrt{3}$) and (3×3) surfaces. Upon further increasing of the coverage, the C₆₀ molecules formed multiple-layer islands on top of the two completely filled C₆₀ layers on the (3×3) surface, as shown in Fig. 3(a). The film structure shows fcc (111) packing with the nearest dis-

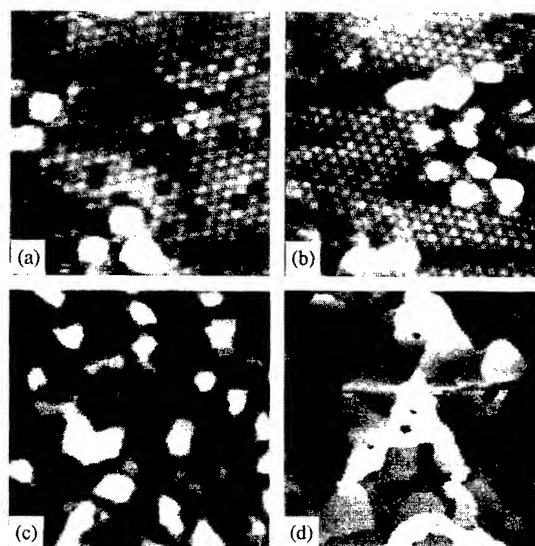


FIG. 4. (a) STM image of the (3×3) surface with 0.02 ML of C₆₀ after annealing to 600 °C (80 Å×80 Å). (b) The ($\sqrt{3}\times\sqrt{3}$) surface with 0.02 ML of C₆₀ after annealing to 600 °C (170 Å×170 Å). (c) The SiC islands grown on the (3×3) surface (480 Å×480 Å) after annealing a single adsorbed C₆₀ layer at 850 °C. (d) Growth of the SiC layer on the (3×3) surface (1000 Å×1000 Å) after annealing the adsorbed C₆₀ layer at 1000 °C.

tance being 10.2 Å. This observation indicates that film growth on the (3×3) surface is dominated by van der Waals interactions. This also suggests that C₆₀-substrate interaction is limited to only the first two layers. Similar results were reported on Si(100) and Ge(100) (2×1) surfaces.^{4,14} Note that the change of contrast at the step edges of the C₆₀ islands in Fig. 3(a) does not represent real physical features; it is the result of a high scan speed. On the other hand, we did not observe formation of a very ordered structure for multilayer coverages on the ($\sqrt{3}\times\sqrt{3}$) surface. This suggests that C₆₀-substrate interaction still influences multilayer growth of C₆₀ on the ($\sqrt{3}\times\sqrt{3}$) surface. Deposition of multilayer C₆₀ molecules led to the growth of three-dimensional clusters. This eventually resulted in the nucleation of large C₆₀ islands. Annealing the C₆₀ film on the (3×3) surface at 400 °C resulted in the desorption of all the C₆₀ molecules, except for the those in the first layer which showed short range ordering [see Fig. 3(b)].

In order to study the decomposition mechanism of the C₆₀ molecule on SiC surfaces, a small amount of C₆₀ was deposited on the surfaces. Then the surfaces were annealed to various temperatures. Annealing the (3×3) surface with 0.02 ML of C₆₀ coverage at temperatures below 600 °C did not change the C₆₀ dimer population or the height of the C₆₀ molecules in the STM images. For temperatures above 600 °C, however, clusters with irregular shapes were formed [see Fig. 4(a)]. This indicates that decomposition of the C₆₀ molecule has occurred. It is also possible that formation of a C₆₀-Si cluster has taken place since there is some depletion of Si immediately surrounding the C₆₀. This observation is consistent with previous studies of the interaction of C₆₀ with the Si(111)-(7×7) surface.¹⁵ For the case of the ($\sqrt{3}\times\sqrt{3}$) surface, annealing the surface with 0.02 ML of C₆₀ coverage

at temperatures up to 1000 °C showed no change of shape of the C₆₀ molecules in the STM images [see Fig. 4(b)]. The height did change, however, to around 3–4 Å in comparison with the 6–8 Å measured for the unannealed C₆₀ molecules. This observation suggests that the C₆₀ molecules did not decompose on the carbon rich (√3×√3) surface. But it is possible that a different bonding configuration at the interface causes the reduction of the height of the C₆₀ molecules in the STM images.

After annealing the adsorbed C₆₀ layer on the (3×3) surface [see Fig. 3(b)] at 850 °C, the C₆₀ reacted with the Si on the surface to form SiC islands as shown in Fig. 4(c). Further annealing to 1000 °C reduced the surface roughness, and resulted in the growth of larger SiC clusters, shown in Fig. 4(d). Atomically resolved images showed that the surface was (√3×√3) reconstructed. In contrast to the behavior observed on the (3×3) surface, annealing a multilayer of C₆₀ adsorbed on the (√3×√3) surface did not result in the formation of any ordered structure. Furthermore, the C₆₀ clusters could not be desorbed by annealing to temperatures as high as 1050 °C. Nevertheless, after annealing the surface in a Si flux at 1000 °C the C₆₀ molecules reacted with the arriving Si to form SiC clusters, with the surface reconstruction being dependent upon the Si coverage. Therefore, growth of SiC can be achieved by alternately supplying Si atoms and C₆₀ molecules to the surface. In this way, growth can be limited to a single atomic layer.

IV. CONCLUSIONS

We have studied the interaction of the C₆₀ molecule with 6H-SiC(0001) (3×3) and (√3×√3) surfaces by STM. The C₆₀ was found to reside at the center of the honeycomb of the C-rich (√3×√3) structure, while on the Si-rich (3×3) surface they resided on top of the bright protrusions. With increasing coverage, the C₆₀ formed triangular or hexagonal clusters on both surfaces. The second layer C₆₀ started to nucleate on the first layer before completion of the first layer, leaving an open space in the first layer. This is explained by the buildup of strain in the C₆₀ clusters on both surfaces, which limited

the growth of the size of the C₆₀ clusters. On the (3×3) surface, starting from the third layer, a film of close-packed fcc (111) with a close nearest neighboring distance of 10.2 Å was observed. On the (√3×√3) surface, the C₆₀ continued to grow in three-dimensional clusters, which eventually led to large C₆₀ islands. The internal structure of the C₆₀ molecule consisted of three stripes. This was observed in the first and second layer C₆₀ on both surfaces. The growth of SiC was achieved by using C₆₀ molecules as the carbon source. Growth can be controlled at the atomic level by surface reconstruction.

ACKNOWLEDGMENTS

This work was partially supported by the Nishina Project (No. 07NP0501) of the Department of Education of Japan. One of the authors (L.L.) thanks the Japan Society for the Promotion of Science (JSPS) for providing him with a post-doctoral fellowship.

- ¹E. A. Rohlfing, D. M. Cox, and A. J. Kaldor, *J. Chem. Phys.* **81**, 3322 (1984).
- ²H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, and R. E. Smalley, *Nature (London)* **318**, 162 (1985).
- ³T. Sakurai, X. D. Wang, Q. K. Xue, Y. Hasegawa, T. Hashizume, and H. Shinohara, *Prog. Surf. Sci.* **51**, 263 (1996).
- ⁴Q. Xue, T. Hashizume, Y. Hasegawa, H. Kamiyama, K. Ohno, Y. Kawazoe, H. Shinohara, and T. Sakurai, *Phys. Rev. Lett.* (submitted).
- ⁵D. Chen, R. Workman, and D. Sarid, *Electron. Lett.* **30**, 1007 (1994).
- ⁶A. V. Haniza, M. Balooch, and M. Moalem, *Surf. Sci.* **317**, L1129 (1994).
- ⁷E. B. Malken and P. Taborek, *J. Appl. Phys.* **78**, 541 (1995).
- ⁸L. Li, Y. Hasehawa, I. S. T. Tsong, and T. Sakurai, *J. Appl. Phys.* **80**, 2524 (1996).
- ⁹L. Li, Y. Hasehawa, I. S. T. Tsong, and T. Sakurai, *J. Phys. (France) IV* **6**, 167 (1996).
- ¹⁰Qikun Xue, T. Ogino, Y. Hasegawa, H. Shinohara, and T. Sakurai, *Phys. Rev. B* **53**, 1985 (1996).
- ¹¹J. Abrefah, D. R. Olander, M. Balooch, and W. Siekhaus, *Appl. Phys. Lett.* **60**, 1313 (1992).
- ¹²T. Sakurai, T. Hashizume, I. Kamiyu, Y. Hasegawa, N. Sano, H. W. Pickering, and A. Sakai, *Prog. Surf. Sci.* **33**, 3 (1990).
- ¹³Y. Guo, N. Karasawa, and W. A. Goddard III, *Nature (London)* **351**, 464 (1991).
- ¹⁴D. Klyachko and D. M. Chen, *Phys. Rev. Lett.* **75**, 3693 (1995).
- ¹⁵M. Balooch and A. V. Hamza, *Appl. Phys. Lett.* **63**, 150 (1993).

Novel C₆₀-DDME complex thin film with electrical bistable properties

M. Ouyang, Z. Q. Xue,^{a)} K. Z. Wang, and Q. D. Wu
Department of Electronics, Peking University, Peking 100871, China

D. Qiang
Department of Chemistry, Peking University, Peking 100871, China

(Received 12 September 1996; accepted 31 March 1997)

A novel C₆₀-DDME complex thin film was prepared by a new modified vacuum deposition technique. Stable and reproducible electrical bistable properties are observed in the C₆₀-DDME thin films. The structure and spectroscopy characteristics of the complex C₆₀-DDME thin film are considerably different from those of both DDME and C₆₀ thin films, as is revealed by high resolution scanning electron microscopy, x-ray diffraction, and ultraviolet-visible absorption spectra. © 1997 American Vacuum Society. [S0734-211X(97)03704-9]

I. INTRODUCTION

Fullerene C₆₀, which was discovered several years ago,¹ is a new allotrope of carbon. Success in the synthesis of macroscopic quantities of C₆₀ (Ref. 2) has aroused intense interests in applications of C₆₀.

So far, electrical switching and memory devices made of inorganic, organic, and metalorganic compounds have been paid considerable attention because of their potential applications in molecular electronic devices such as high speed switching and memory elements.³ Since Potember *et al.* reported stable and reproducible switching phenomena between two resistant states induced by an applied electric field in the charge transfer complex thin films of Ag-TCNQ and Cu-TCNQ,^{4,5} in-depth studies of electronic and optical properties of such devices have been carried out extensively.^{6,7} For example, data writing and erasing in Ag-TCNQ and Cu-TCNQ thin films realized by scanning tunnel microscopy (STM) have been reported.^{8,9} This would create ultrahigh density information storage up to 10¹² bits/cm² possible and greatly enrich this field of research.¹⁰ In addition to C₆₀'s extraordinary properties such as superconductivity or photoconductivity caused by alkali metal and other materials doping,¹¹⁻¹⁴ applying C₆₀ to the research field of information storage has also been implemented.^{15,16}

Recently, a new organic material, 1,1-dicyano-2-(4-dimethylaminophenyl)ethene (DDME), was synthesized in our laboratory. The molecular structure of DDME is shown in Fig. 1. It is a polynitrile π acceptor, and the donor-acceptor system of Ag-DDME complex thin films is proven to possess electrical bistable properties.¹⁷ Here, the structural characteristics and the reproducible bistable switching from a high- to a low-impedance state of a novel C₆₀-DDME complex thin film are presented.

II. EXPERIMENTS

The C₆₀-DDME complex thin films were prepared by following a modified vacuum deposition technique. A schematic diagram of the vacuum evaporation jig is shown in

Fig. 2. The vacuum pumping system for preparing the thin films consisted of a rotary pump and a diffusion pump. There was a heated U tube placed vertically in the vacuum chamber. The inner diameter of the U tube was 1.5 cm, the U tube was 4.5 high, and the hand-to-hand distance was 2.5 cm. A sample holder was located at one end of the U tube. The distance between the sample holder and the end of the U tube was 1.5 cm. Two sources for simultaneous deposition of DDME and C₆₀ were set just under the other end of the U tube.

The following preparation steps were all made at a vacuum of 1.0×10^{-3} Pa. First, the Au thin film of about 100 nm was evaporated onto a glass substrate as a bottom electrode. Then the Au film on the glass substrate was put onto the sample holder. The U tube was heated to 130 °C by a tungsten filament surrounding the U tube for about 15 min and was maintained at 130 °C during deposition. The aim of this procedure was to make the evaporating molecules attain thermal equilibrium with the heated wall of the U tube by random collisions, resulting in a uniform nucleation of the molecules on the substrate, and to prevent deposition from taking place on the wall of the U tube instead of on the substrate. Afterward, the crucibles, loaded with DDME powders and C₆₀ powders, respectively, were heated simultaneously. In order to mix them homogeneously, it was important to control the deposition rate and the deposition time of both the DDME and the C₆₀ properly. Once the C₆₀-DDME complex had grown on the Au film to the desired thickness, a corresponding complex thin film was formed. Finally, we deposited another Au thin film of about 100 nm as a top electrode. Thus, the sandwichlike system consisting of layers of Au/C₆₀-DDME/Au was complete.

The dc *I-V* curves of the C₆₀-DDME thin film were measured at room temperature in air with a 10 k Ω resistor as the series load. High resolution scanning electron microscopy (HRSEM) photos were taken with an Amray 1910 field emission microscope. X-ray diffraction (XRD) measurements were made on a Rigaku D/max-3B diffractometer with a Cu *K α* source. Ultraviolet-visible (UV-vis) absorption spectra of the films on the quartz substrates were measured

^{a)} Author to whom correspondence should be addressed.

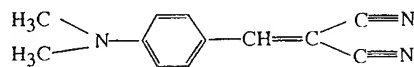


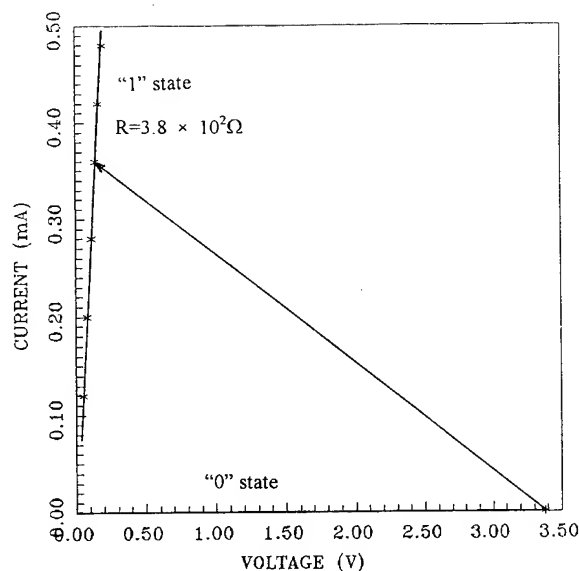
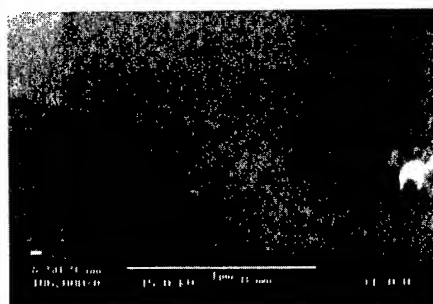
FIG. 1. The molecular structure of DDME.

on a Shimadzu 240 spectrophotometer, using blank quartz substrate as the reference.

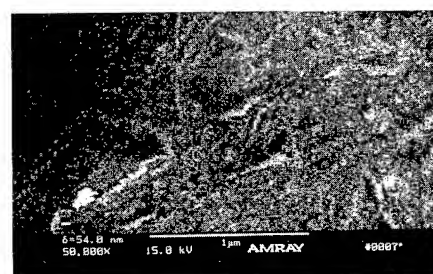
III. ANALYSIS OF THE RESULTS

Figure 3 is the typical I - V curve of C₆₀-DDME ($\sim 0.8 \mu\text{m}$ thick). The complex thin film is initially in a high-resistance state when applied voltage is lower than 3.38 V (field strength $\sim 4.2 \times 10^4 \text{ V/cm}$), and the resistance is about $10^9 \Omega$. At a threshold voltage of 3.38 V the complex thin film switches abruptly from high resistance ("0" state) to low resistance ("1" state). The resistance of this conductive state is about 380Ω . The resistance difference between the 0 state and the 1 state is about seven orders (10^7). No preliminary forming such as a thermal treatment is required for the complex thin film to be able to switch. In the 1 state, the voltage can be increased or decreased to 0 V without switching to the 0 state. The memory time of the 1 state in this typical device is about 38 h. It is worth noting that the memory time of the C₆₀-DDME complex thin films that we prepared is longer than that of similar devices such as C₆₀-BDCP and C₆₀-BDCB complex thin films.^{15,16} When heated in ambient atmosphere at 70 °C for 1.5 h, the complex thin film in the 1 state can switch back to the 0 state. Such an electrical bistable switching characteristic is stable and reproducible.

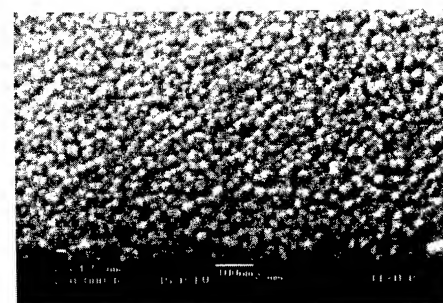
The structural analyses of C₆₀-DDME complex thin films demonstrate that C₆₀-DDME is a new phase that is different from both pure C₆₀ thin films and neutral DDME thin films. Figure 4 shows the HRSEM images of a pure C₆₀ thin film, a DDME thin film, and a C₆₀-DDME thin film (these thin films were all prepared by our new method mentioned above). From Fig. 4(c), one can see that there are small clusters with almost the same diameter (less than 20 nm) distributed in the C₆₀-DDME complex thin film, but Figs. 4(a) and 4(b) show that the pure C₆₀ thin film and the pure DDME thin film possess a smooth surface. The comparison of XRD patterns of C₆₀ thin film, DDME thin film, and C₆₀-DDME thin film in Fig. 5 also supports this conclusion. The XRD

FIG. 3. The typical I - V curve of C₆₀-DDME.

(a)



(b)



(c)

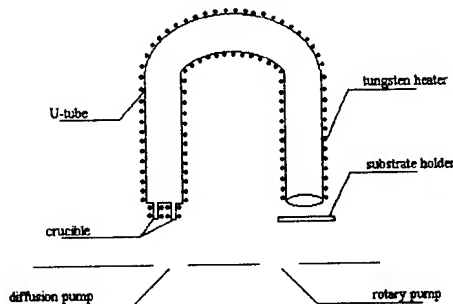
FIG. 4. HRSEM images of (a) a pure C₆₀ thin film, (b) a pure DDME thin film, and (c) a C₆₀-DDME complex thin film.

FIG. 2. Schematic diagram of the vacuum evaporation jig.

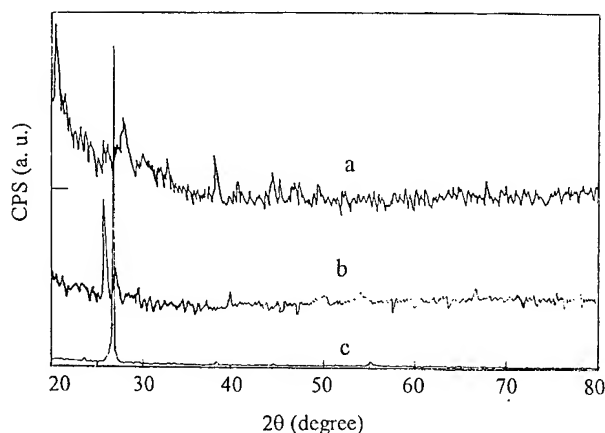


FIG. 5. XRD patterns of (a) a pure C₆₀ thin film, (b) a pure DDME thin film, and (c) a C₆₀-DDME complex thin film.

pattern of the C₆₀-DDME film shows significant differences from those of both the DDME and C₆₀ films alone. There are diffraction peaks at $2\theta = 25.9^\circ$, 26.8° , 28.2° , and 38.3° in the DDME film and at $2\theta = 20.8^\circ$, 25.6° , 26.0° , 27.9° , 28.2° , 32.9° , 38.3° , 44.6° , 45.4° , and 67.9° in the C₆₀ film. However, in the C₆₀-DDME these characteristic peaks of C₆₀ and DDME thin films almost disappear and only a strong peak at $2\theta = 26.5^\circ$ exists.

The C₆₀-DDME complex thin films show very different spectral features in comparison with the DDME thin films and the C₆₀ thin films. There are five absorption peaks at 227, 295, 363, 567, and 601 nm in the DDME thin films [see Fig. 6(b)], and in the C₆₀ thin films [see Fig. 6(a)] there are absorption peaks at 221, 268, 348, 568, and 618 nm. The C₆₀-DDME complexes [see Fig. 6(c)] exhibit a wide band centered at 445 nm and two weak peaks at 560 and 590 nm. Unfortunately, since the crystal structure of the C₆₀-DDME complex thin film is not yet known exactly, it would be difficult to give a more exact interpretation of the UV-vis spectrum of the C₆₀-DDME thin films.

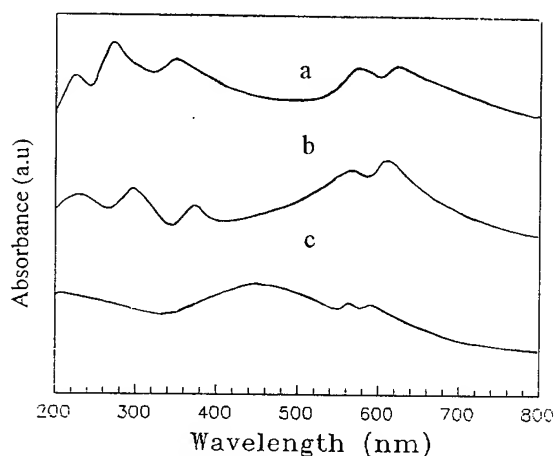


FIG. 6. UV-vis absorption spectra of (a) a pure C₆₀ thin film, (b) a pure DDME thin film, and (c) a C₆₀-DDME complex thin film.

Three mechanisms, the formation of conductive filaments,¹⁸ the charge transfer mechanism,^{4,5} and the changes in molecular or crystal structures,¹⁰ have been suggested to interpret these electrical switching phenomena. Nevertheless, the transition caused by the formation of conductive filaments is unstable and usually has appreciable physical damage on the surface. The electrical bistable property of the C₆₀-DDME complex thin films are stable and reproducible, however, and no physical damage of the surface of C₆₀-DDME films was observed under a microscope. Therefore the formation of conductive filaments in the present device may be ruled out. In addition, because the thin films of both pure DDME and C₆₀ are highly insulating, the switching behavior of C₆₀-DDME films cannot be ascribed to the charge transfer mechanism. So the switching phenomenon of C₆₀-DDME complex thin films that is observed is most probably due to the change in molecular or crystal structures, the same result as that in some previous work.^{15,16}

IV. CONCLUSIONS

A novel organic material, 1,1-dicyano-2-(4-dimethylaminophenyl)ethene (DDME), was synthesized and C₆₀-DDME complex thin films were prepared by the modified vacuum deposition technique. A stable and reproducible electrical bistable property of C₆₀-DDME complex thin films was found, and the structural and spectral characteristics of thin films were studied. Advanced research related to C₆₀-DDME complex thin films, such as electrical information storage by STM and optical information storage, is now in progress.

¹H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, and R. E. Smalley, *Nature (London)* **318**, 162 (1985).

²Z. Gu, J. Qian, X. Zhou, Y. Wu, X. Zhu, S. Feng, and Z. Gan, *J. Phys. Chem.* **95**, 9615 (1996).

³R. S. Potember, *Molecular Electronic Devices*, edited by F. L. Carter (New York, 1992), p. 73.

⁴R. S. Potember, T. O. Poehler, and D. O. Cowan, *Appl. Phys. Lett.* **34**, 405 (1979).

⁵R. S. Potember, T. O. Poehler, and A. Rappa, *J. Am. Chem. Soc.* **103**, 3659 (1980).

⁶R. S. Potember, T. O. Poehler, and R. C. Benson, *Appl. Phys. Lett.* **41**, 548 (1982).

⁷Z. Y. Hua and G. R. Chen, *Vacuum* **43**, 1019 (1992).

⁸M. Matsumoto, Y. Nishio, H. Tachibana, T. Nakamura, Y. Kawabata, H. Samura, and T. Nagamura, *Chem. Lett.* **234**, 1021 (1991).

⁹S. Yamaguchi, C. A. Viands, and R. S. Potember, *J. Vac. Sci. Technol. B* **9**, 1129 (1991).

¹⁰K. Z. Wang, H. X. Zhang, D. W. Wang, and Z. Q. Xue, *Vac. Sci. Technol. (China)* **16**, 277 (1996).

¹¹A. F. Hebard, M. J. Rosseinsky, R. C. Haddon, D. W. Murphy, S. H. Glarum, T. T. M. Palstra, A. P. Ramirez, and A. R. Kortan, *Nature (London)* **350**, 600 (1991).

¹²M. J. Rosseinsky, A. P. Ramirez, S. H. Glarum, D. W. Murphy, R. C. Haddon, A. F. Hebard, T. T. M. Palstra, A. R. Kortan, S. M. Zahurak, and A. V. Makhija, *Phys. Rev. Lett.* **66**, 2830 (1991).

¹³Y. Wang, *J. Phys. Chem.* **96**, 764 (1992).

¹⁴Y. Wang, *Nature (London)* **356**, 585 (1992).

¹⁵M. Ouyang, K. Z. Wang, H. X. Zhang, Z. Q. Xue, C. H. Huang, and D. Qiang, *Appl. Phys. Lett.* **68**, 2441 (1996).

¹⁶K. Z. Wang, M. Ouyang, H. X. Zhang, Z. Q. Xue, C. H. Huang, and D. Qiang, *Acta Sci. Nat. Univ. Pekinen.* **32**, 628 (1996).

¹⁷K. Z. Wang, Z. Q. Xue, M. Ouyang, D. W. Wang, H. X. Zhang, and C. H. Huang, *Chem. Phys. Lett.* **243**, 217 (1995).

¹⁸H. Hamann, *Int. J. Electron.* **73**, 1039 (1992).

Si- and C-rich structure of the 6H-SiC(0001) surface

L. Li,^{a)} Y. Hasegawa, and T. Sakurai

Institute for Materials Research, Tohoku University, Sendai 980-77, Japan

(Received 10 March 1997; accepted 10 March 1997)

The reconstructions of the 6H-SiC(0001) surface under both Si-rich and C-rich conditions were studied using field ion-scanning tunneling microscopy (FI-STM). The sample was cleaned by *in situ* Si beam etching at 900–1000 °C. The as-cleaned surface showed a ($\sqrt{3}\times\sqrt{3}$) structure. The Si-rich phases were produced by annealing the sample in a Si flux. With increasing Si concentration, (2×2), (2 $\sqrt{3}\times 6\sqrt{3}$), (3×3), and (7×7) reconstructions were observed. Reaction of the Si-rich phases with C₂H₂ molecules at 1050 °C resulted in the formation of a C-rich surface, which exhibited a (2×2)/(6×6) reconstruction. A structure model for ($\sqrt{3}\times\sqrt{3}$) reconstruction is proposed, and possible applications of using the surface reconstruction to selectively grow SiC polytype is discussed. © 1997 American Vacuum Society. [S0734-211X(97)01304-8]

I. INTRODUCTION

Silicon carbide (SiC) is a wide band gap semiconductor with potential applications for short wavelength optoelectronic, high temperature, radiation resistant, and high-power electronic devices. SiC crystallizes in either a cubic or a hexagonal form, and exhibits polytypism, which is derived from different stacking sequences of Si/C bilayers along the closest packed direction. If the stacking sequence is ABCABC..., one obtains a single SiC polytype, which is referred to as 3C or β -SiC. If the stacking sequence is ABAB..., then a variety of hexagonal polytypes is obtained. The most common hexagonal polytypes are 6H and 4H, where the 6 and 4 refer to the number of planes in the periodic stacking sequence. Combination of both stacking sequences yields a wide range of ordered, larger period hexagonal or rhombohedral structures. These are referred to as α -SiC. Despite identical stoichiometry SiC polytypes differ in their electronic energy band gap by as much as 1 eV. This opens up interesting perspectives for development of multilayer structures of different polytypes, creating new materials with novel properties.¹ Thus, understanding the mechanism of controlling the SiC polytype growth is of practical and theoretical interest.² Molecular beam epitaxy (MBE) has been used to grow SiC at about 1000 °C.^{3,4} It was found that the surface structure of the substrate plays a critical role in controlling the growth kinetics. A recent study suggested that atomic layer epitaxy (ALE) is a likely way to achieve the growth of definite SiC polytypes.⁵ Therefore, detailed analysis of the atomic structure of the 6H-SiC surface is needed to clarify the conditions necessary to obtain a specific polytype of SiC.

II. EXPERIMENT

The experiments were carried out in an ultrahigh vacuum (UHV) chamber containing a scanning tunneling microscope (STM) equipped with a low energy electron diffraction (LEED) and a field ion microscope (FIM) which was used to

monitor and fabricate the STM tip. Details of the FI-STM experimental setup have been discussed elsewhere.⁶ The samples were *n* type ($n=2\times 10^{18}\text{ cm}^{-3}$) 6H-SiC(0001) single crystals supplied by Cree Research. The surface was cleaned *in situ* by Si molecule beam etching at 900–1000 °C. Temperature was measured using an optical pyrometer with an emissivity setting of 0.53. The Si beam was produced by heating a small Si slab to 1100–1150 °C positioned 10 cm away from sample. The Si flux was estimated to be in the range of 10^{14} – $10^{15}\text{ cm}^{-2}\text{ s}^{-1}$. The carbon source was acetylene (C₂H₂), which was introduced into the chamber through a leaking valve. All STM images were taken with sample bias of –2.0 V, unless specified otherwise.

III. RESULTS AND DISCUSSION

After exposing the SiC(0001) surface to the Si flux at 950 °C for 5 min, the surface was clean and exhibited large terraces of up to 500 Å. The majority of steps had a height of about 7 Å, which is close to three Si/C bilayer high (7.6 Å). On the terraces, high resolution dual bias images (Fig. 1) showed a honeycomb structure in filled state (a) and a hexagonal structure in empty state (b) with height contrast reversal. The measured unit cell spacing was $5.4 \pm 0.2\text{ Å}$, in good agreement with the expected value of 5.3 Å for a ($\sqrt{3}\times\sqrt{3}$) reconstruction, and corrugation of the surface was typically 0.2 Å. This structure has been reported in previous studies,^{7–9} and a model with an Si adatom at the *T*₄ site has

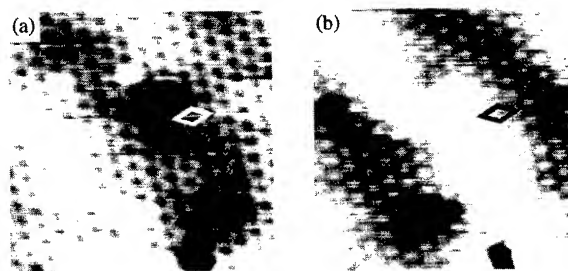


FIG. 1. (a) Filled state (–1.2 V) and (b) empty state (+1.2 V) image of the ($\sqrt{3}\times\sqrt{3}$) reconstruction (70 Å×70 Å).

^{a)}Present address: Chemical Engineering Department, University of California, Los Angeles, CA 90095; Electronic mail: lian@seas.ucla.edu

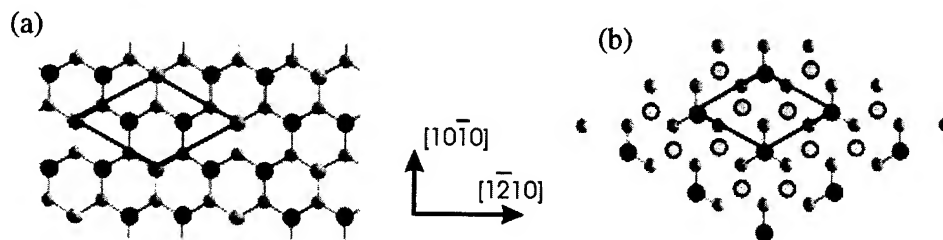


FIG. 2. (a) Top view of the surface vacancy model. Large and small balls are the Si and C atoms of the Si/C bilayer, respectively. Dark balls are the C atoms at the surface. (b) Top view of the Si-C alloy model. Top layer Si is shown by dark balls. The substitution C atoms are shown in circles.

been suggested. Recent first-principles total-energy calculations also support the Si- T_4 adatom model. The calculated electronic structure shows a metallic and half-filled surface state, which suggests that the STM images should be similar for tunneling in and out of the surface.¹⁰ However, our STM result shown in Fig. 1 is inconsistent with the Si- T_4 adatom model since the STM images in the filled state and empty state show a height contrast reversal. This Si- T_4 model is also inconsistent with a recent photoemission study which showed a surface state with semiconducting occupation for the $(\sqrt{3}\times\sqrt{3})$ reconstruction.¹¹

Based on our previous STM results, a surface vacancy model for the $(\sqrt{3}\times\sqrt{3})$ structure was proposed.¹² Figure 2 shows the surface vacancy model and another possible surface Si-C alloy model for the $(\sqrt{3}\times\sqrt{3})$ reconstruction. In the surface vacancy model, a layer of carbon with one-third of its atoms missing on top of the Si-C bilayer creates a pattern of vacancies which has the $(\sqrt{3}\times\sqrt{3})$ geometry. The C adatoms reside on the atop sites, and the dangling bond of the Si atom at the center of the vacancy is unsaturated. In the surface Si-C alloy model, 2/3 of the surface Si atoms are substituted by C atoms in a manner that creates the $(\sqrt{3}\times\sqrt{3})$ geometry. In both models, we assume 2/3 of ML of carbon on the surface although the crystallographic orientation for the surface is Si-terminated. The fact that this $(\sqrt{3}\times\sqrt{3})$ is formed at about 1000 °C suggests that Si depletion on the surface is likely. Previous x-ray photoemission spectroscopy (XPS) and Auger electron spectroscopy (AES) studies have indicated that carbon segregation can occur at temperature as low as 900 K.¹³ Simple electronegativity considerations indicate that charge transfer is mainly from Si to C. As a result of the charge transfer, it can be expected that C atoms are likely to be revealed in filled state STM images and Si atoms to be revealed in empty state STM images. However, we should note that the relationship between STM images and surface geometric structure is complicated. In our proposed models, hybridization of the dangling bonds of the surface Si and C atoms can be expected. Due to a lack of calculations of the energy levels of the Si and C dangling bonds, comparison of them to the conduction band minimum and valence band maximum of 6H-SiC is impossible. The proposed models for the $(\sqrt{3}\times\sqrt{3})$ structure are therefore tentative, and further studies are needed.

When the $(\sqrt{3}\times\sqrt{3})$ surface was exposed to the Si beam, four more distinct surface phases have been found as shown in Fig. 3. The first phase formed is shown in Fig. 3(a), after

deposition of Si on the $(\sqrt{3}\times\sqrt{3})$ surface at 850°C for about 1 min. The surface showed local (2×2) ordering. Figure 3(b) shows the surface after further deposition of Si on the (2×2) surface, and the structure was determined to be $(2\sqrt{3}\times 6\sqrt{3})$ reconstruction. The atomic arrangement of the $(2\sqrt{3}\times 6\sqrt{3})$ structure can be seen to form two types of rows, one type containing four maxima and the other containing two maxima. Deposition of more Si on the $(2\sqrt{3}\times 6\sqrt{3})$ produced a hexagonal array of protrusions in both filled and empty state image with corrugation of 2.5 Å that formed a (3×3) reconstruction [Fig. 3(c)]. It is apparent that our STM images of the (3×3) reconstruction are once again inconsistent with a model proposed based on the Si(111)-(7×7) dimer-adatom-stacking fault (DAS) model in the previous LEED and energy loss spectroscopy (ELS) studies,⁷ since the very prominent corner holes of the DAS (3×3) structure are absent in both filled state and empty state STM images. Eventually, the DAS (7×7) structure was produced by deposition of Si on the (3×3) surface as shown in Fig. 3(d). Accompanying the (7×7) reconstruction, the step structure also changed from relatively straight when the surface was $(\sqrt{3}\times\sqrt{3})$ reconstructed to triangular shaped. The step edges were

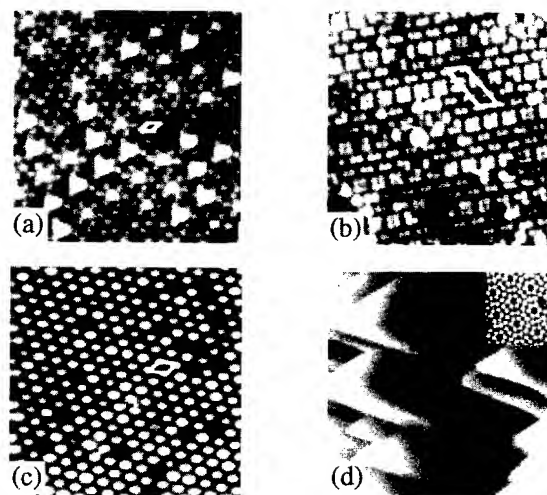


FIG. 3. Sequence of STM images of the 6H-SiC(0001) surface, showing the Si-rich structures in the order of increasing Si concentration: (a) The (2×2) reconstruction ($120\text{ Å}\times 120\text{ Å}$); (b) the $(2\sqrt{3}\times 6\sqrt{3})$ reconstruction ($118\text{ Å}\times 118\text{ Å}$); (c) the (3×3) reconstruction ($160\text{ Å}\times 160\text{ Å}$); (d) the (7×7) reconstruction ($1200\text{ Å}\times 1200\text{ Å}$). The insert shows an empty state image (+2 V) of the (7×7) structure.

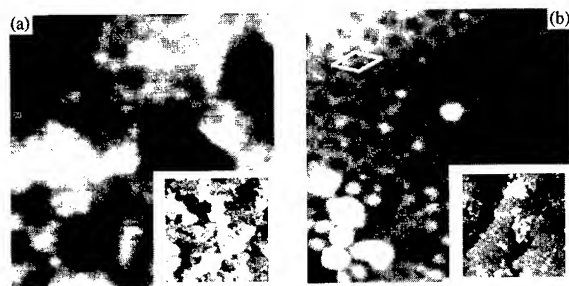


FIG. 4. STM images of the C-rich structure formed after interaction of the Si-rich phases with C_2H_2 molecule at 1050 °C. (a) The (2×2) structure after 30 L of C_2H_2 exposure to the (3×3) surface ($90 \text{ \AA} \times 90 \text{ \AA}$). The sample bias is +2.0 V. The image size for the inset is $1200 \text{ \AA} \times 1200 \text{ \AA}$. (b) The empty state image (+1.6 V) of the (6×6) structure ($170 \text{ \AA} \times 170 \text{ \AA}$). The inset shows a large scale image of the surface ($1200 \text{ \AA} \times 1200 \text{ \AA}$).

oriented along $\langle 1120 \rangle$ directions. The change of surface morphology likely resulted from the large lattice mismatch between Si and SiC (20% for Si and β -SiC). Annealing the surface at about 1000 °C resulted in Si desorption, which led to successive transformation among these Si-rich phases. The order of phase transformation is from (7×7) to (3×3) to $(2\sqrt{3} \times \sqrt{3})$ to (2×2) . The surface Si concentration decreases in the same order.

The carbon-rich phases were prepared by exposing the Si-rich phases to C_2H_2 at 1000–1050 °C. Carbon clusters were formed on the surface after exposing the (3×3) surface to 30 L of C_2H_2 at 1050 °C as shown in Fig. 4(a). Atomically resolved STM images showed local (2×2) ordering. At saturated coverage (100 L), about 1 ML of carbon was deposited on the surface [Fig. 4(b)], and the surface showed a (6×6) structure. Similar structures were observed after reaction of the C_2H_2 molecule with other Si-rich phases. The Si-rich phases can be produced by exposing the (2×2) and (6×6) surface to the Si flux. Therefore, epitaxial growth of SiC has been achieved by alternate supply of Si and C_2H_2 molecules to the surface. However, it is difficult to determine the polytype by imaging the surface structure only.

Figure 5 shows the phase diagram summarizing the observed reconstructions of the 6H-SiC(0001) surface as a function of the surface Si and C composition.

IV. CONCLUSIONS

We have shown that *in situ* Si beam etching is very effective to remove the oxide of the 6H-SiC(0001) surface. The 6H-SiC(0001) surface exhibits a variety of reconstructions, which can be used to regulate the growth process at an atomic level. Furthermore, since the reconstructed surface has fewer dangling bonds, and the symmetry of the reconstructed surface is different from the (1×1) surface, the growth energetics and kinetics are expected to be different.

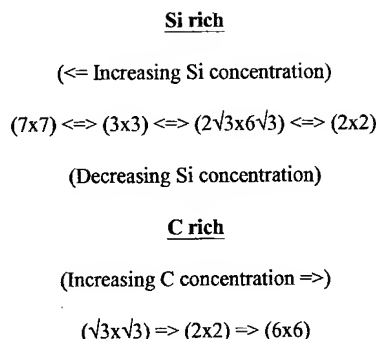


FIG. 5. Phase diagram of the 6H-SiC(0001) surface as a function of surface Si and C composition.

As in the case of the $(2/3 \times 6/3)$ surface [see Fig. 2(b)], the rows will provide an easier diffusion path for adatoms compared with diffusion across the rows. Indeed, study has shown that SiC films grown on the 6H-SiC(0001)- (3×3) surface resulted in the step flow, and subsequently the SiC films were 6H polytype. While films grown on the (1×1) surface resulted in 2D nucleation, the SiC films were β -SiC polytype.⁴ The present study provides detailed structural information of the 6H-SiC(0001) surface as a function of surface Si and C composition. The results will help us understand the surface chemistry and mechanism of SiC polytype growth by MBE.

ACKNOWLEDGMENTS

This work was partially supported by the Nishina Project (No. 07NP0501) of the Department of Education of Japan. One of the authors (L. L.) thanks the Japan Society for the Promotion of Science (JSPS) for providing him with a post-doctoral fellowship.

¹F. Bechstedt and P. Käckell, Phys. Rev. Lett. **75**, 2180 (1995).

²V. Heine, C. Cheng, and R. J. Needs, J. Am. Ceram. Soc. **74**, 2630 (1991).

³S. Tanaka, R. S. Scott, and R. F. Davis, Appl. Phys. Lett. **65**, 2851 (1994).

⁴A. Fissel, B. Schroter, and W. Richter, Appl. Phys. Lett. **66**, 3182 (1995).

⁵A. Fissel, U. Kaiser, K. Pfennighaus, B. Schroter, and W. Richter, Appl. Phys. Lett. **68**, 1204 (1996).

⁶T. Sakurai, T. Hashizume, I. Kamiyu, Y. Hasegawa, N. Sano, H. W. Pickering, and A. Sakai, Prog. Surf. Sci. **33**, 3 (1990).

⁷R. Kaplan, Surf. Sci. **215**, 111 (1989).

⁸F. Owman and P. Martensson, Surf. Sci. **330**, L639 (1995).

⁹L. Li and I. S. T. Tsong, Surf. Sci. **351**, 141 (1996).

¹⁰J. E. Northrup and J. Neugebauer, Phys. Rev. B **52**, R17001 (1995).

¹¹L. I. Johansson, F. Owman, and P. Martensson, Surf. Sci. **360**, L478 (1996).

¹²L. Li, Y. Hasegawa, I. S. T. Tsong, and T. Sakurai, J. Appl. Phys. **80**, 2524 (1996).

¹³L. Muehlhoff, W. J. Choyke, M. J. Bozack, and John T. Yates, Jr., J. Appl. Phys. **60**, 2842 (1986).

Adsorption geometries of 1,9-decadiene on Si(111) 7×7 studied by scanning tunneling microscopy

D. Shachal and Y. Manassen

Department of Chemical Physics, Weizmann Institute of Science, Rehovot 76100, Israel

(Received 12 September 1996; accepted 31 March 1997)

The scanning tunneling microscope has been used to characterize the adsorption sites of 1,9-decadiene on Si(111) 7×7. The molecule which reacts with each of its two ends, appears as pairs of reacted adatoms on the atomically resolved surface. The size of the molecule permits adsorption with the two ends of the molecule situated either at the same half of the unit cell or in a bridging geometry with one end situated in the faulted half and the other in the unfaulted half. Three types of bridging pairs are found. The results found are compared with that of two other dienes: 1,7-octadiene and 1,13-tetradecadiene. © 1997 American Vacuum Society. [S0734-211X(97)04504-6]

I. INTRODUCTION

Understanding the chemistry of adsorbates on surfaces has been a subject for numerous studies using various surface-sensitive techniques. The scanning tunneling microscope (STM) has proven to be a good tool in studying the chemistry of small molecules on the atomic level.¹⁻⁶ In these systems, the only factor affecting the reactivity is the relative affinity between the molecules reacting group and the different available sites. In a previous article,⁷ we have taken a step towards understanding more complex systems, bifunctional molecules. For these molecules, the adsorption depends both on the nature of the reacting groups and the steric matching between these reacting groups and the available surface sites. Three diene molecules, differing by the chain length separating the two double bonds at each end of the molecules, were imaged on a Si(111) 7×7. This surface is interesting because of the richness and diversity of its available sites, the dangling bonds. The unit cell can be divided into two subunits, the faulted and the unfaulted halves. Their different reactivity, proven by experiment and theory,^{4,8,9} is attributed to a higher density of states in the faulted half. In each half, the available sites are the rest atoms and the center and corner adatoms.

The binding of each end of a diene molecule can be understood on the basis of the chemistry of ethylene. The molecule adsorbs nondissociatively at room temperature. High resolution electron energy loss spectroscopy (EELS) studies showed that the molecule is bi- σ bonded: the double bond breaks to form two Si-C bonds.¹⁰ Piancastelli *et al.* using tunneling spectroscopy proposed that ethylene is bonded to a pair of adjacent Si adatom rest atoms.⁴ The aliphatic chain, a saturated hydrocarbon, is not expected to react with the surface. Since STM images at a sample bias of +2 V reveal only the adatoms, a diene molecule is expected to appear as a pair of reacted adatoms separated by a typical length related to the molecule size.

In the previous study,⁷ we have shown the feasibility of imaging diene molecules. For 1,6-heptadiene and 1,7-octadiene, the adsorption was to sites restricted to one half of the unit cell. The molecules appeared as reacted pairs of

either center-center adatoms or center-corner adatoms. The faulted half was found to be more reactive. The size of a longer molecule, 1,13-tetradecadiene, imposes reaction with one end of the molecule situated in the faulted half and the other end in the unfaulted half. These pairs will be referred as bridging pairs. In this article, we report an intermediate case, the adsorption of 1,9-decadiene. The molecule was chosen because its size permits adsorption either to sites located at one half or to sites in neighboring halves. The goal of this study is to characterize the adsorption geometries of the molecule.

II. EXPERIMENT

The experiments were carried out under UHV (base pressure 2×10^{-10} Torr). The *n*-type Si(111) wafer was outgassed overnight and flashed to 1000 °C in order to remove the oxide layer. After a few cycles of freezing and pumping, the molecule 1,9-decadiene (Fluka 97%) was dosed through a leak valve from a gas manifold. During the dosing, the gas manifold was differentially pumped to maintain the ratio between the molecule vapor pressure and the background pressure as large as possible. Typically, imaging began 3 h after the flash. The images shown were taken at a sample bias of +2 V and tunneling current of 1 nA.

III. RESULTS AND DISCUSSION

Regions of a surface exposed to 1,9-decadiene are seen in Fig. 1. The surface exhibits dark protrusions grouped into pairs. The pairs are either situated in the same half or bridged between neighboring halves. Images of 1,7-octadiene and 1,13-tetradecadiene are presented in Fig. 2 for comparison. The center-center and center-corner adatoms pair found for 1,7-octadiene are also the dominant pairs found for 1,9-decadiene. The new characteristic found is the bridging pairs. They constitute about a quarter of the pairs found (153 out of 573) while for 1,7-octadiene, their number was too small for evaluation. Three types of bridging pairs are found. They are either composed of nearest neighbor center-center adatoms (molecule 1 in Fig. 1) or nearest neighbor center-

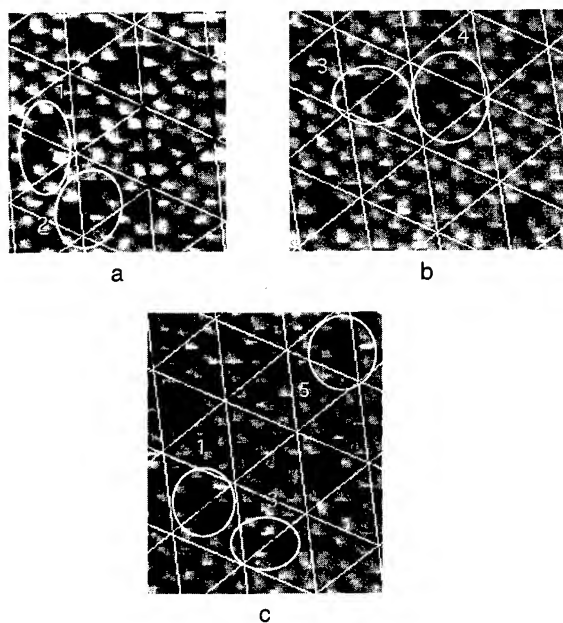


FIG. 1. (a)–(c) Adsorbed molecules of 1,9-decadiene on Si(111) 7×7 . Lines intersecting the unit cells are drawn for clarity. The different combinations of pairs are outlined: The bridging pairs composed of center-center adatoms (1), center faulted-corner unfaulted adatoms (2), corner faulted-center unfaulted adatoms (3), and pairs restricted to a single half: center-center adatoms (4) and center-corner adatoms (5).

corner adatoms (molecules 2 and 3 in Fig. 1). Since the two halves are not equivalent, a center-corner bridging pair can be either center-faulted corner unfaulted (molecule 2), or corner-faulted center unfaulted (molecule 3). One can see that the bridging pairs found for 1,13-tetradecadiene [Fig. 2(b)], composed of next nearest neighbor center ada-

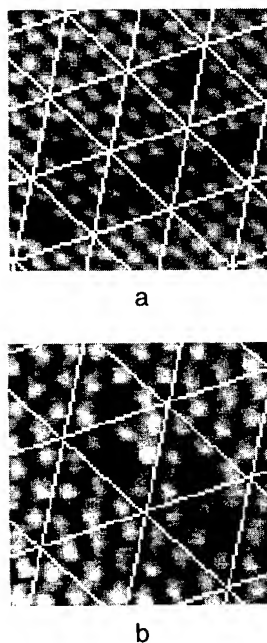


FIG. 2. (a) Adsorbed molecules of 1,7-octadiene on Si(111) 7×7 . (b) Adsorbed molecules of 1,13-tetradecadiene on Si(111) 7×7 .

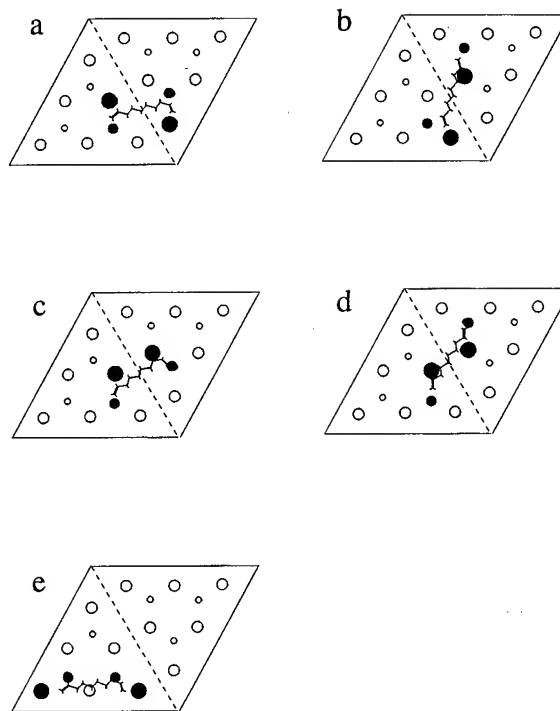


FIG. 3. The possible sites on the surface giving rise to adsorption of 1,9-decadiene for a center-corner bridging pair (a) and (b) and a center-center bridging pair (c) and (d). The adatoms (large circles) and rest atoms (small circles) that react are solid filled. A molecule is schematically placed on the unit cell. Both are shown on the same scale. (e) A molecule placed with its ends above two corner adatoms.

toms pair, reflect a larger distance between the molecule ends as expected. As a routine, we also imaged the filled states of the surface, to determine the position of the faulted and unfaulted halves. At a negative bias, the reacted sites also appear as dark protrusions. This is consistent with the contrast found for ethylene by Piancastelli *et al.*⁴

Although the images reveal only the reacted adatoms, it is assumed that each double bond reacts with a pair of neighboring adatom rest atoms. The possible adsorption geometries for the bridging pairs are sketched in Fig. 3. The presence of two rest atoms next to each center adatom gives rise to two possible geometries [(a) and (b), (c) and (d) in Fig. 3] for each center-corner and center-center pair, respectively. These geometries are indistinguishable from the images. An intriguing fact is the absence of a pair composed of two reacted corner adatoms, both for the bridging geometry and for molecules situated in the same half. Inspecting the combination of sites for adsorption within a single half, one can see that the combination of two corner adatoms fits best the molecular length as shown in Fig. 3(e). Two factors may explain these observations. First, an electrostatic repulsion between the dangling bonds situated between the two corner adatoms and the aliphatic chain may disfavor this geometry. Another possibility is the existence of a precursor state for adsorption associated only with the center adatom and not with the corner adatom. Piancastelli *et al.*⁴ have suggested that ethylene adsorbs to Si(111) 7×7 via a mobile precursor state. For the diene molecules studied however, the precursor

state may be immobile due to the size of the molecules and the presence of two reacting groups. In such a case, not all the combinations of sites are available in particular the corner-corner adatom pair.

IV. SUMMARY

The study of 1,9-decadiene adsorption reflects another aspect of the role of the chain length on the reactivity of bi-functional molecules. The matching between the length of the molecule and the surface sites offers two routes for reaction: adsorption with the two ends in one half or adsorption in a bridging geometry to both halves. Each of these routes was demonstrated in a previous study: the size of 1,6-heptadiene and 1,7-octadiene restricts the molecules to adsorb via the first route while the size of 1,13-tetradecadiene forces reaction via a bridging geometry. In the future, we plan to investigate whether the competition between these different channels can be controlled for example by heating the surface or photoinjecting carriers to change the distribution between the two halves.

ACKNOWLEDGMENTS

This work was supported by the Minerva Foundation, Munich, Germany and the Basic Research Foundation administered by the Israeli Academy of Sciences and Humanities. One of the authors (Y.M.) is an incumbent of the Lilian and George Lyttle Career Development Chair.

¹R. Wolkow and Ph. Avouris, *Phys. Rev. Lett.* **60**, 1049 (1988).

²Ph. Avouris, I. W. Lyo, F. Bozco, and E. Kaxiras, *J. Vac. Sci. Technol. A* **8**, 3405 (1990).

³Ph. Avouris and I. W. Lyo, *Surf. Sci.* **242**, 1 (1991).

⁴M. N. Piancastelli, N. Motta, A. Sgarlata, A. Balzarotti, and M. De Crescenzi, *Phys. Rev. B* **48**, 17 892 (1993).

⁵A. J. Mayne, A. R. Avery, J. Knall, T. S. Jones, G. A. D. Briggs, and W. H. Weinberg, *Surf. Sci.* **284**, 247 (1993).

⁶M. Chander, Y. Z. Li, J. C. Patria, and J. H. Weaver, *Phys. Rev. B* **48**, 2493 (1993).

⁷D. Shachal, Y. Manassen, and E. Ter-Ovanesyan, *Phys. Rev. B* **55**, 9367 (1997).

⁸J. Yoshinubu, D. Fukushi, M. Uda, E. Nomura, and M. Aono, *Phys. Rev. B* **46**, 9520 (1992).

⁹K. D. Brommer, M. Galvan, A. Dal Pino, Jr., and J. D. Joannopoulos, *Surf. Sci.* **314**, 57 (1994).

¹⁰J. Yoshinobu, H. Tsuda, M. Onchi, and M. Nishijima, *Solid State Commun.* **60**, 801 (1986).

Study on surface and interface structures of nanocrystalline silicon by scanning tunneling microscopy

Juning Gao, Haiqiang Yang, Ning Liu, Dongxia Shi, Yueshan Jiang, Zengquan Xue, and Shijin Pang

Laboratory of Vacuum Physics, Center for Condensed Matter Physics, Chinese Academy of Sciences, P.O. Box 2724, Beijing 100080, China

Yuliang He

Amorphous Physics Research Laboratory, Beijing University of Aeronautics and Astronautics, Beijing 100083, China

(Received 8 September 1996; accepted 13 May 1997)

The scanning tunneling microscope has been employed to study the morphology, atomic surface structures, and grain interface of hydrogenated nanocrystalline silicon (nc -Si:H) before and after hydrofluoric acid (HF) etching. It was found: (1) The nc -Si:H films were composed of many different sizes of grains and these grains were composed of many finer grains. (2) There were line structures on the surface of the fine grains and loop structures at the grain boundaries without HF treatment. After etching, two more structures were observed: loop structures on the surface of fine grains and spider bonding structures besides the interface of fine grains. (3) The loop structures found at the grain boundaries was larger and more irregular than those on the grain surfaces. Line structures were similar to crystal silicon, but the distance between lines was enlarged. Considering the experimental results, a discussion was made about the formation mechanism of these atomic structures. © 1997 American Vacuum Society. [S0734-211X(97)12904-3]

I. INTRODUCTION

In the last decade, hydrogenated nanocrystalline silicon (nc -Si:H) films were widely used in microelectronic devices, such as thin-film transistors.¹ Recently, some researchers reported the photoluminescence of nc -Si:H at room temperature.^{2,3} This may bring about some new applications of nc -Si:H. The morphology, grain size, and grain boundary of nc -Si:H are important characteristics for the applications of nc -Si:H thin films in microelectronics. The properties of nc -Si:H, such as electrical conductivity, are strongly influenced by the grain size and grain boundary,⁴⁻⁶ so the research of the microstructures of nc -Si:H is essential for its applications. In the past, most research work of nc -Si:H was done by transmission electron microscopy, scanning electron microscopy (SEM), x-ray diffraction, and Raman spectroscopy. Recently, scanning tunneling microscopy (STM) has been used in the studies of nc -Si:H.^{4,7} The high resolution of STM made it possible to observe surface structures from micron to atomic resolution in real space and various environments.⁸ However, these air-exposed surfaces are difficult to observe because of a native oxide layer. Such an oxide layer can be removed by chemical etching in hydrofluoric acid (HF). Nakagawa *et al.* have reported the observing of Si(100) and Si(111) atomic images in air after HF treated.⁹ Niwa, Iwasaki, and Hasegawa also have reported the observation of quite clear STM images for HF-dipped surfaces of Si(100).¹⁰ However, what will happen on the surface and interface of nc -Si:H after HF etching has not been investigated.

In this article, STM was used to study the morphology, grain size, and grain boundary before and after HF etching,

especially to observe the atomic structures of the surface and grain boundary.

II. EXPERIMENT

The nc -Si:H films were deposited on a glass substrate in a capacitance system of a conventional plasma-enhanced chemical vapor deposition with hydrogen (H_2) diluted silane (SiH_4) as the reactants activated by rf and dc double power sources. During the deposition process, the substrate temperature was at 300 °C, the rf power was 0.44 w/cm², and the total pressure of the reactive gases was 1.3–1.5 Torr. The details of the experimental parameters have been described elsewhere.¹¹

The STM used in the experiment is a home-made CSTM-9000 operated in air at room temperature. The tip was from a mechanically sharpened platinum–iridium alloy wire. The STM images were taken in the constant current mode. To remove the oxidized layer, the samples were dipped into 10% HF solution for about 20 s. With the prolongation of scanning, the STM images degraded, this was due to some surface reaction caused by the tunneling current.^{7,9} When the scanning area was larger than 25 nm×25 nm, the bias voltage was about 2 V and the constant tunneling current was set to about 0.5 nA. To study the atomic structure of the surface, the bias voltage and the constant tunneling current were set to 1.5 V and 1.5 nA, respectively.

III. RESULTS AND DISCUSSION

Figure 1 is a STM image of the nc -Si:H film. The scanning area is 50 nm×50 nm. There is a larger grain among several grains with different sizes in the image. The grain sizes measured by STM varied from 10 to 100 nm. There are



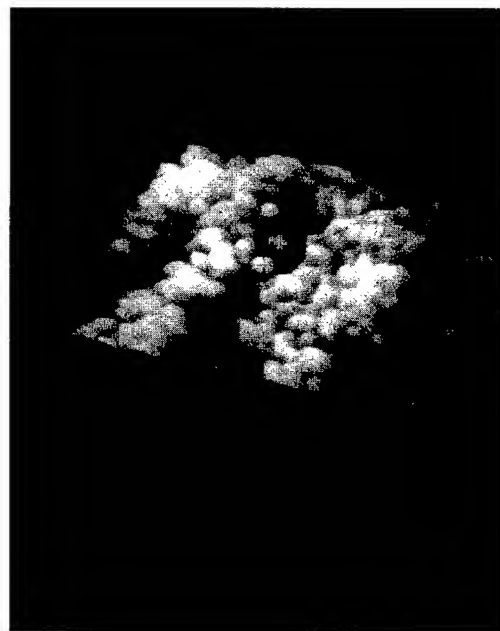
FIG. 1. STM image of *nc*-Si:H. The scale is 50 nm×50 nm.

lots of fine grains on the surface of the larger grain, and the size of the fine grain is about several nanometers. Figure 2 shows the fine grains on a larger grain. The scanning area is 25 nm×25 nm. The existence of fine grains increases the area of interfaces, so the hydrogen content increases due to the existence of a lot of Si-H bonds on the surfaces of the grains. This result can explain the high content of hydrogen in *nc*-Si:H in our previous study.⁶ The existence of Si-H bonds prevented the surface from being oxidized heavily and made it possible to observe the surface and interface without HF etching.

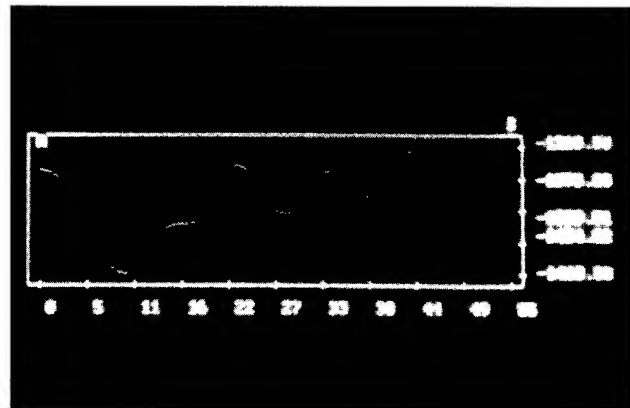
Figures 3, 4, and 5 show the atomic structures of fine grains and their interfaces obtained after HF etching. The



FIG. 2. STM image of fine grains. The scale is 25 nm×25 nm.



(a)



(b)

FIG. 3. (a) Two kinds of loop structures after HF etching. The scale is 4.22 nm×3.14 nm. (b) Cross section of (a). A: loop structure in the interface, and B: loop structure on the surface of the fine grain.

scanning area of Fig. 3(a) is 4.22 nm×3.14 nm. In Fig. 3 we can find two kinds of loop structures marked A and B, respectively (A: loop structure on the interface; B: loop structure on the surface). A line was drawn to cross these two kinds of loop structures and Fig. 3(b) shows the cross section along the line. Considering the morphology and the cross section, we find that A is larger and more irregular than B. Figure 4 shows another kind of atomic structure: a line structure (marked C). The scanning area was 2.61 nm×2.25 nm. The distance between two lines is about 0.45 nm, which is different from that of the crystal faces of crystal silicon. Figure 5 shows a common and interesting structure: a spider bonding structure. Also, this structure can be considered as a loop structure and many line structures instead of a new one. There is a center in the spider bonding structure, which is made up of an atom, a vacancy, or a cluster. The other atoms surrounding the center arrange themselves as a spider web.

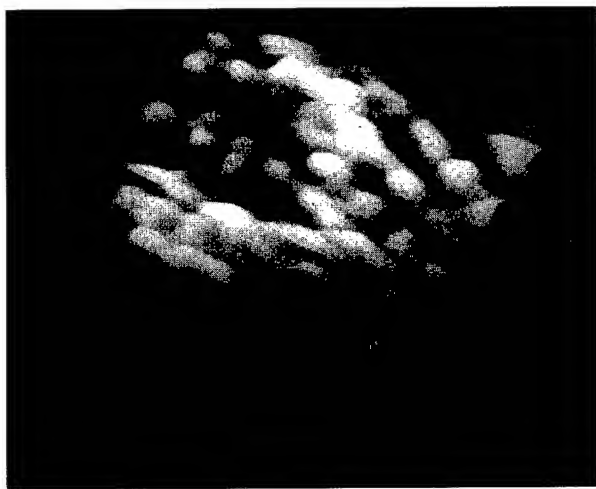


FIG. 4. Line structure (marked C) after HF etching. The scale is $2.61 \text{ nm} \times 2.25 \text{ nm}$.

Etching may play a very important role in the formation of all these atomic structures. Figures 6 and 7 show STM images without HF treatment. In Fig. 6, we can find the line structure (marked D). Figure 7 shows a loop structure. Considering these different orientations around the loop, we may draw the conclusion that this loop structure is on the interface of the fine grains. Contrasted with the atomic structure after HF etching, loop structure B and the spider bonding structure are not found. This result is consistent with our previous work.^{6,11} Furthermore, before HF etching, it is dif-

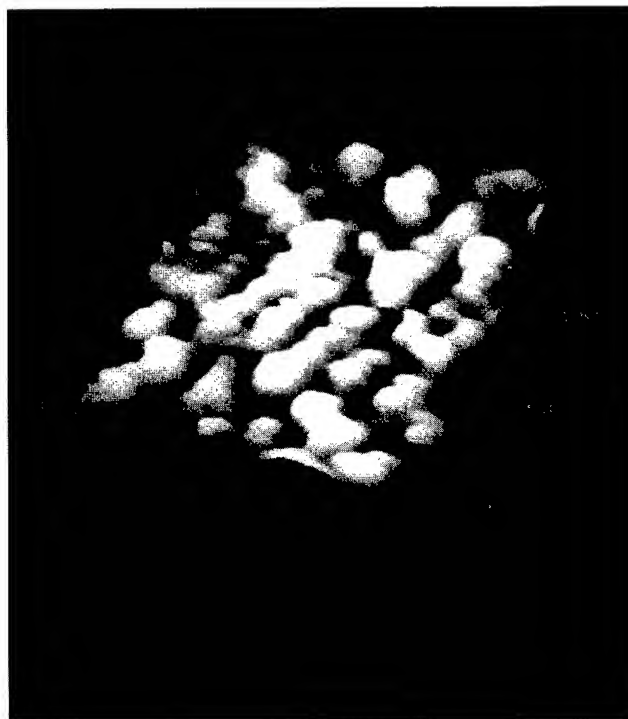


FIG. 6. Line structure in the surface without HF treatment. The scale is $1.96 \text{ nm} \times 1.82 \text{ nm}$.

ficult to tell apart the single atoms from that after the HF treatment, correspondingly, even if they are a similar distance between the rows. The degradation of resolution may be caused by oxidation. On the other hand, the loop structure is very similar to loop structure A, which is obtained after HF etching.

It is well known that an air-exposed *nc*-Si:H surface always has a native oxide layer. This layer can restrict the resolution of the images to a nanometer level.⁹ While removing the oxide layer on the surface by HF etching, the bond of the silicon and the oxygen was broken. Then, dangling bonds of the silicon were produced. At the same time, many ions of hydrogen exist in the solution. In the process of the combination of atoms, the silicon atoms can combine with each other or with the hydrogen ion. On the surface of the fine grains, due to the various distances between rows of the crystal lattice, two kinds of structures, line structure and loop structure, can be formed, respectively. When the distance between the rows is close enough, to reduce the dangling bond, silicon atoms in one row would combine with the nearest atoms in the same row and those in the nearest rows. This situation would form the loop structures on the grain. To the contrary, when the distance between the rows is large, silicon atoms between the rows cannot combine. Therefore, they will combine with the hydrogen ion in the solution. So, we can find line structures on the surface. If the distance is neither too large to form the line structure nor too small to form the loop structure independently, then the loop structure and line structure will exist disordered and nonsystematic. This random collocation can be found elsewhere, i.e., in Fig. 3. In

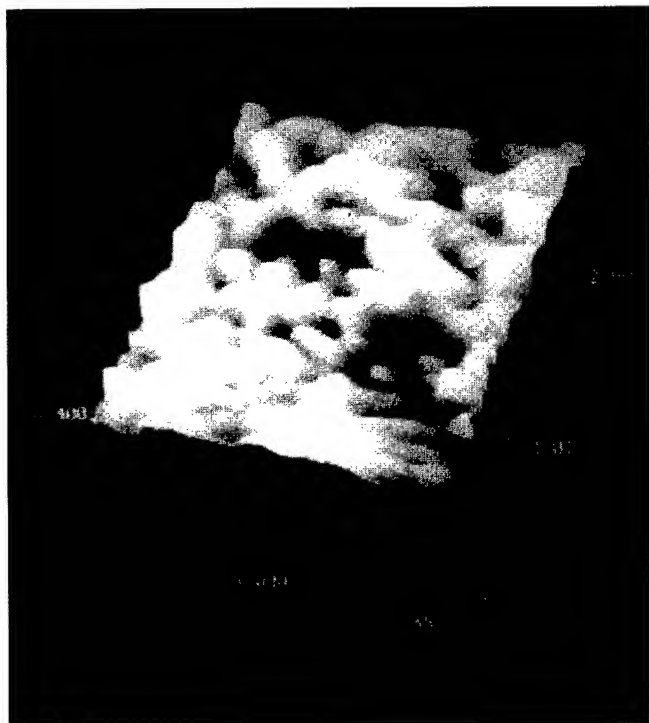


FIG. 5. Spider bonding structure beside the interface of fine grains after HF etching, $2.31 \text{ nm} \times 2.31 \text{ nm}$.

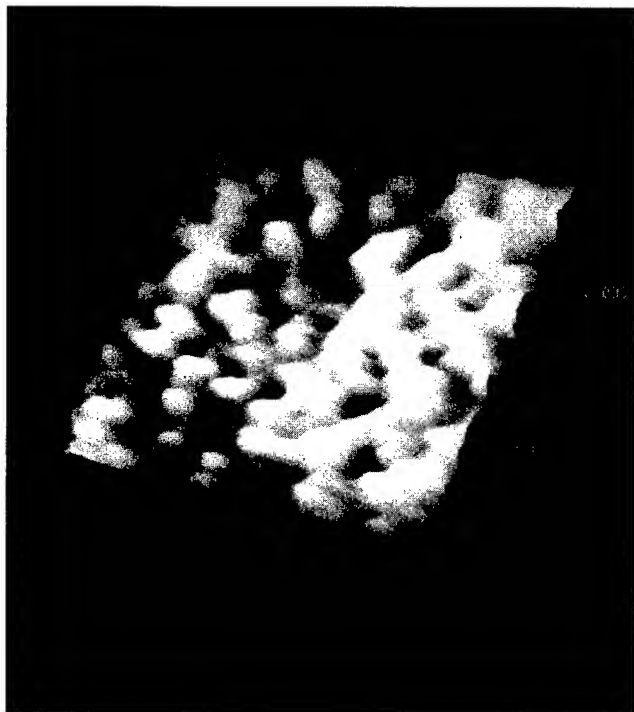


FIG. 7. Loop structure in the interface of fine grains before HF etching. The scale is $2.19 \text{ nm} \times 2.11 \text{ nm}$.

comparison with the line structure, the spider bonding structure is formed near the interface. While the center atoms combine together, the other atoms have combined with the hydrogen ion as a result of the distance being enlarged. This is why the spider bonding structure can be considered as loop and line structures. Yet, we must verify that spider bonding structures are formed only beside the interface of the fine grains.

At the site of the interface, where several finer grains meet together, the loop structure on the interface can be formed. This is not because of the HF etching but because of the procedure of deposition. At the beginning of the film formation, many atoms were located and diffused on the surface of the substrate randomly. When some atoms met together, they formed a small cluster. Then the cluster as a whole grew and formed a fine grain in the end. While several fine grains met together, the larger loop structures were

formed. These structures are similar to the interface model proposed by Gleiter.⁵ While in etching by HF, this kind of loop structure may keep its initial form because the atoms will combine with the hydrogen ion only.

IV. CONCLUSION

The surface structures of *nc*-Si:H were studied by STM. It was found that the films were composed of many different sized grains and the grains were composed of many fine grains. Atomic images of the fine grains and their boundaries were successfully observed before and after HF etching. Three kinds of atomic structures were investigated: (1) Loop structures: After HF etching, this kind of structure was observed both on the surface of the grain and at the grain boundaries. The loop structures found at the grain boundaries were larger and more irregular than those on the surface of the grain. Consistent with Fig. 6 without HF treatment, only the loop structures on the interface can be found. We thought that the loop structures on the interface were caused by the congregation of fine grains and the loop structures on the surface were caused by HF etching. (2) Line structures: This kind of structure was similar to crystal silicon, but the distance between the lines was different. (3) Spider bonding structures: This kind of structure can only be formed beside the grain boundaries after HF etching. So, we can draw a conclusion that the etching may play a very important role in the formation process of all these atomic structures.

¹Hong Joo Lim, Bong Yeol Ryu, and Jin Jang, *Appl. Phys. Lett.* **66**, 2888 (1995).

²Xiang-na Liu, Xiao-wei Wu, Xi-mao Bao, and Yu-liang He, *Appl. Phys. Lett.* **64**, 220 (1994).

³Hideki Tamura, Markus Ruckschloss, Thomas Wirschem, and Stan Veprek, *Appl. Phys. Lett.* **65**, 1537 (1994).

⁴J. K. Gimzewski, A. Humbert, D. W. Pohl, and S. Veprek, *Surf. Sci.* **168**, 795 (1986).

⁵H. Gleiter and Saarbrücken, *Europhys. News* **20**, 130 (1989).

⁶Yuling He, Chenzhong Yin, Guangxu Cheng, Luchun Wang, Xiangna Liu, and G. Y. Hu, *J. Appl. Phys.* **75**, 797 (1994).

⁷I. Tanaka, F. Osaka, T. Kato, Y. Katayama, S. Muramatsu, and T. Shimada, *Appl. Phys. Lett.* **54**, 427 (1989).

⁸G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, *Phys. Rev. Lett.* **49**, 57 (1982).

⁹Y. Nakagawa, A. Ishitani, T. Takahagi, H. Kuroda, H. Tokumoto, M. Ono, and K. Kajimura, *J. Vac. Sci. Technol. A* **8**, 262 (1990).

¹⁰Masaaki Niwa, Hiroshi Iwasaki, and Shigehiko Hasegawa, *J. Vac. Sci. Technol. A* **8**, 266 (1990).

¹¹Y. L. He, X. Liu, Z. C. Wang, G. X. Cheng, L. C. Wang, and S. D. Yu, *Sci. China A* **36**, 248 (1993).

Diffusional attractions between voids on a Si(111)7×7 surface

E. Ter Ovanessian, Y. Manassen, N. Ramesh Rao, and Z. Olami

Department of Chemical Physics, The Weizmann Institute of Science, Rehovot 76100, Israel

(Received 12 September 1996; accepted 31 March 1997)

Ostwald ripening or coarsening is a term which describes the growth of large domains at the expense of small ones in the final stages of any phase separation process. Several theories were developed in order to explain this process. All of them are based on the Lifshitz, Slyozov, and Wagner (LSW) model which predicts an increase in the average radius of domains $\bar{r}(t) \propto t^{1/3}$ and a reduction of the number density of domains as a function of time $N(t) \propto t^{-1}$. The LSW model assumes static and circular domains, which are distributed at random. These assumptions were found to be incorrect by many experiments. And more advanced theories were necessary to explain these observations. In our work, we study the coarsening of voids formed in a Si(111)7×7 surface covered by 0.8 bilayers (BL). Besides the expected increase in the diameter of voids and the decrease in the number density, very strong attractions and rapid motion of voids towards each other were clearly observed. The result of the attraction between the voids is coagulation of voids and an increase in the spatial correlations between voids, as the coarsening process progresses, as we indeed observed. This is contradictory to the "ideal gas" picture which is the basis of the LSW model. Basically, this can be explained by a concentration gradient which develops between two voids of different size and the growth of voids in the direction of the larger concentration. However, this model cannot explain the large diffusional motions observed in our experiments. It is proposed that an amplification mechanism exists, whereby the arrival of a diffusing vacancy to the boundary of the void causes the release of several adatoms from the boundary. This process causes a much larger motion than expected by a single vacancy absorption. This picture, which is consistent with the structure of the steps, demonstrates how the microscopic details might have a significant affect on the global coarsening process. © 1997 American Vacuum Society. [S0734-211X(97)04203-0]

I. INTRODUCTION

A phase separation process begins with nucleation and is followed by growth of small droplets of one phase dispersed in the second one. At later stages, the process is usually diffusion controlled and is characterized by the growth of large domains at the expense of small ones. This process is known as Ostwald ripening¹ or coarsening. It is dominated by the diffusion of material from smaller domains to larger ones. The driving force for the growth is the size dependent stability of the domains.

Such coarsening processes had been observed in many different systems. Examples for it in three dimensions are the aging of a binary alloy² and the phase separation of a solid-liquid mixture, formed by heating a solid alloy above the solidus temperature;³ Another example is the coarsening of diblock copolymers.^{4,5} A diblock copolymer is a linear chain molecule consisting of two subchains labeled a and b. A repulsion between a and b will lead to a partial phase separation and to the formation of microdomains. Another case, which is relevant to this work is the process of sintering.^{6,7} After the formation of the porous solid, most of the pores are gradually diminishing and finally disappear or migrate to the surface. The driving force here, also, is the tendency of the solid to minimize the curvature of the surface of the pore.

A porous solid is usually characterized by a small amount of pores in the solid. In this case, the minority phases are the pores and the diffusion is a diffusion of vacancies (rather than atoms or molecules). This picture, also relevant in our

case, is entirely valid in describing coarsening of a porous solid.

Two dimensional coarsening is also observed in an extremely large number of systems. A typical example is the coarsening of solid domains surrounded by a metal in a thin film,⁸ coarsening of a monomolecular film made from a mixture of two immiscible liquid phases of two amphiphilic molecules,⁹ and the ordering of an adsorbed layer on a metallic surface.¹⁰ Several scanning probe microscopy studies which are related to coarsening were performed.¹¹⁻¹⁴

The basic theory of Ostwald ripening was developed by Lifshitz, Slyozov,¹⁵ and Wagner¹⁶ (LSW). According to this theory, there is a time dependent critical domain size $r^*(t)$, which is the boundary between shrinking and growing domains. Domains smaller than $r^*(t)$ will shrink and domains larger than $r^*(t)$ will grow. According to this theory, $r^*(t) = \bar{r}(t)$, where \bar{r} is the average domain radius. The average radius grows in time as $\bar{r} \propto t^\alpha$, where $\alpha=0.33$. Namely, the average domain volume increases linearly with time. Since the total fraction of the material in the minority phase is constant, the number of minority phase domains decreases according to $N(t) = t^{-1}$ (in three dimensions).

The LSW theory is a mean field theory dependent on a single parameter c_m which is the (mean field) concentration of the diffusing atom or molecule (or vacancies, in the case of pores). The small domains have a larger concentration γ/r (where γ is the surface tension) than c_m at their interface.

As a result, atoms diffuse away from smaller domains into the large ones. Smaller domains will disappear while larger ones will grow. Over the course of time, c_m will diminish and \bar{r} will grow. Though the LSW theory captures the essential ingredients of the coarsening dynamics, it has questionable assumptions and thus might be regarded as a good starting mean field theory for other, more realistic theories. First, it is clear that the mean field assumptions will break very rapidly when the fraction of the minority phase is not negligible. It is obvious that a small domain which is a supply of diffusing atoms to the neighboring ones, will make them larger, since the local concentration is larger than c_m . The opposite is also true: A large domain will help in making its neighbors smaller. Thus the growth of a domain cannot be determined only by the global average c_m , but rather by the local concentration, which is dependent on the local arrangement of the domains. A larger than the average domain may diminish if surrounded by other large ones. A smaller than average domain may grow if surrounded by other small ones.

This leads to an expected size anticorrelation which is contradictory to one of the basic LSW assumptions, namely that the domains are distributed randomly. Many theoretical works were performed in order to take this anticorrelation effect into account, both in three and two dimensions.^{17–20} These improved theories show that the $t^{1/3}$ power law for the increase of the average (and critical) radius of the domain is still valid, both in three and two dimensions. However, the distribution function is modified. While the LSW distribution is not symmetric, the distributions observed from these theories are broader and more symmetric.

Two other questionable assumptions of the LSW theory is that the domain positions are fixed in space, and that they have a perfect circular shape. Experiments, and also a more advanced theory, show that this is not necessarily true.^{3,21} In the systems studied,²¹ nonspherical shapes of the domains were observed though the interfacial energy is completely isotropic. This is a result of diffusional interactions between the domains. A simulation which allows the domains to change shape as a result of a nonuniform concentration of the interface showed that deviations from a spherical shape are possible. When this deviation is strong, there will be a strong tendency to return to a circular shape. Thus there is a strong size effect here since the driving force for reaction is proportional to γ/r . So a large domain will stay at a nonspherical shape for a longer time than smaller domains, which will have a larger tendency to remain circular.

An even larger deviation from the basic LSW assumptions was observed for the assumption of a fixed, time independent position of the domains. It was found that the strong diffusional interactions can induce significant motion of the center of masses of the domains. The reduced interdomain distances will cause the development of time dependent correlations.²¹

The origin of this motion, also observed at a relatively small fraction of the minority phase is, bluntly, the growth of the domain at one side, and the shrinking at the other side, as a result of nonuniform diffusion fluxes due to the presence of

neighboring domains. The motion is induced by diffusional interactions. These interactions between the coarsening particles are long range, due to the slow decay in the diffusion field of a single domain.

A final assumption in the LSW theory (which is also entirely incorrect in our case) is the assumption that the only way of transporting particles between domains is by interdomain particle diffusion. However, in many cases, where there is a finite fraction of the minority phase, a significant probability for collision and coagulation of the domains is observed. Phenomenological theories that include effects of coagulation (not in the LSW case) were developed,²² leading, again, to a symmetrization and broadening of the distribution function of the radii of the domains.

At this point, it is important to emphasize that the ripening process is not always diffusion controlled. In cases where attachment or detachment into the domain is a rate determining process, a power law of $t^{1/2}$ is expected for the average domain radius rather than $t^{1/3}$ (Ref. 23). In this case, all the diffusion related processes that are mentioned above will not affect the ripening dynamics.

In this article, a scanning tunnel microscope (STM) study of two dimensional coarsening of voids on a Si(111)7×7 surface is described. The process is initiated on a surface with a large number of randomly distributed small voids which is prepared by the crystallization of a submonolayer amorphous silicon.^{24,25} Vacancy diffusion generates a two dimensional realization of a sintering process. In the experiment, the sample is heated consecutively to accelerate the vacancy migration. At room temperature, the diffusion of vacancies is negligible and the process is frozen, so we can use the relatively slow STM technique to observe images which represent the process at the end of each heating period.

Thus we can get detailed information of the two dimensional coarsening process of the voids. This enables us to observe how many of the assumptions of the naive mean field coarsening theory fail in this system. We observe void movement, coagulation, and correlations between clusters and nonspherical void shapes. We suggest microscopic and global mechanisms which are probably responsible for those effects.

II. EXPERIMENT

The experiments were performed with a custom built STM with a manual lever demagnifier, and custom built electronics and computer software. The Si(111)7×7 sample was degreased by rinsing in trichloroethylene, with subsequent rinsing in acetone, methanol, and deionized water. After overnight degassing, the samples were cleaned by few short annealings to 1100 °C. The duration of the annealing was about 10 s, to keep the pressure in the 10^{-10} Torr range. Annealing to 900 °C for 5 min was followed by a slow cooling (1–2 °C/s) and low-energy electron diffraction (LEED) inspection to verify the quality of the 7×7 reconstruction.



FIG. 1. A STM image of a Si(111) substrate after deposition of 0.8 BL of silicon followed by annealing to 460 °C for 10 min. Note that there is a step in the lower right side of the image. The darker voids there, are in a lower terrace. Area: 60×60 nm².

A thin amorphous silicon film was deposited on the Si(111)7×7 surface from an electron beam evaporator. The film thickness in the experiments reported here was 0.8 BL (1.6 ML). The thickness and the deposition rate (about 0.1 BL/s) were controlled by a quartz crystal monitor. After each period of heating (usually to 600°), the STM was used to image the surface. The temperature was measured with an infrared pyrometer which was calibrated *in situ* with a thermocouple. The thermocouple, mounted on a small linear feedthrough in the STM chamber, was pressed on the sample during the temperature measurement. The accuracy of the temperature measurement is estimated to be 50 °C. All images were obtained with a 2 V sample bias and a tunneling current of 1 nA. The STM measurements were performed at room temperature. Waiting one hour after each heating period was necessary to let the sample and the sample holder cool enough, such that no thermal drift was observed during scanning.

III. RESULTS

After deposition of 0.8 BL of silicon on a Si(111) surface, STM imaging revealed a structureless surface without any sign of the 7×7 reconstruction. After further annealing the sample to 480 °C, a surface covered with small islands connected into a continuous network was observed. Numerous voids of very irregular shape are visible in the layer. Figure 1 shows a crystallized layer with many types of surface reconstructions: The (3×3), (5×5), (7×7), (9×9), (2×2), and ($\sqrt{3} \times \sqrt{3}$)R30 reconstructions were all observed in the images. Also, the atomic structure of the uncovered part of the surface was converted from the 7×7 reconstruction to the mixed structure characteristic to the crystallized layer.



FIG. 2. A STM image of a Si(111)7×7 substrate after deposition of 0.8 BL of silicon at room temperature, followed by annealing to 600 °C for 8 min. Area: 56×56 nm².

This is an example of the starting configuration of the surface from which the coarsening process begins. The two phases (the voids and the upper bilayer) are still not separated, and are mixed to a large extent. The voids have a very irregular shape which has a large circumference to area ratio, and therefore, have also a large (2D) surface tension. The surfaces, both the covered parts and the uncovered ones, have many vacancies, thus, the initial concentration of these diffusing elements is high. The surface structure consists of many different reconstructions, distributed at random. Therefore, in the initial stages of the experiments, the diffusion constant of vacancies is isotropic. As will be shown later, since the coarsening process occurs parallel with ordering of the surface, at later stages, some anisotropy is introduced in the diffusion process.

After annealing of this sample to 600 °C for 8 min, the structure of an observed surface looks completely different (Fig. 2). First, the voids have an almost circular shape, and have a small circumference to area ratio. In regions where the voids are not circular, a straight boundary along the corner holes is observed. Second the surface looks much more ordered and a much smaller amount of vacancies are seen in the image. The surface reconstruction is largely a 7×7 superlattice, and only a small part of it, near the void, has a 5×5 structure. This means, that now the diffusion constant is no longer isotropic, and the diffusion is preferred in directions parallel to the 7×7 corner holes. A grain boundary (marked by the white arrow) is seen in the image to connect between the void and a nearby step. A grain boundary network that connects between the different voids was found in many images. It should be recalled that this process is a two dimensional realization of the process of sintering. In sintering, the (3D) grain boundaries were shown to play a significant role in the process as diffusion sinks and diffusion pipes.⁷ This is so, since the bulk solid, the diffusion constant

is much larger along the grain boundary than anywhere else.

In the surface, this effect is probably less dominant, but the fact that grain boundaries are, indeed, visible between many voids, shows that they have some role in the observed anisotropy which will be discussed later.

An additional important fact is the existence of steps. Their presence complicates the analysis of the results. In order to observe meaningful data, it is necessary to minimize the number of steps as much as possible by using a surface as close as possible to the ideal (111) plane. In addition, it is necessary to ignore the voids that are too close to the step. It was shown in a previous publication²⁵ on a more vicinal surface that the initial step edge is completely straight, while as the coarsening process progresses, the step edge loses its straight shape, and starts to coarsen during the process. The structure around the step also provides some important information regarding the mechanism of diffusion on the surface.

There are two basic atomic mechanisms of transport in this system. One possibility, which was mentioned already is the diffusion of vacancies between the voids. The second possibility is the diffusion of adatoms: These adatoms are released first from the step and diffuse on the surface—between the voids. In contrast to the diffusion of adatoms, diffusion of vacancies between neighboring sites requires overcoming the barriers to diffusion that exist due to the presence of neighboring atoms in the *same* adlayer. Therefore, the activation energy for surface diffusion of adatoms is lower than for vacancy migration. So one would expect this mechanism to be dominant near a step.

The observed images at the steps show that the second transport mechanism is dominant near the steps. In all the step edges, there is an asymmetric distribution of voids. Four examples are shown in Fig. 3. In all the steps seen in the figure, the upper terrace has several voids, while a comparable area in the lower terrace has no voids. This is a result of the dominance of the adatom release from the steps. The adatoms which are released from the step move into the voids near the step and cause their disappearance.

One should note that the basic mechanism in this process is a random walk of the adatoms on the surface. Since the probability to return to the step and to be absorbed on it again is large, a strong density gradient will be created near the steps so the coarsening process far away from the edge will not be affected.

We note that this adatom emission and adsorption is a main mechanism for the roughening of the step during the process. A more detailed description will be given elsewhere. One should note that a release of an adatom from the boundaries of the void to the surface is a process with a much higher energy threshold than a simple release from a step. The main transport mechanism in the void coarsening is either vacancy diffusion or adatom jumps. Both processes are much slower than the release from the step.

In one case, a set of six experiments was performed on the same sample. An amorphous layer of a thickness of 0.8 BL was deposited on a Si(111)7×7 surface. The surface was annealed to 550 °C for 1 min to create an initial distribution

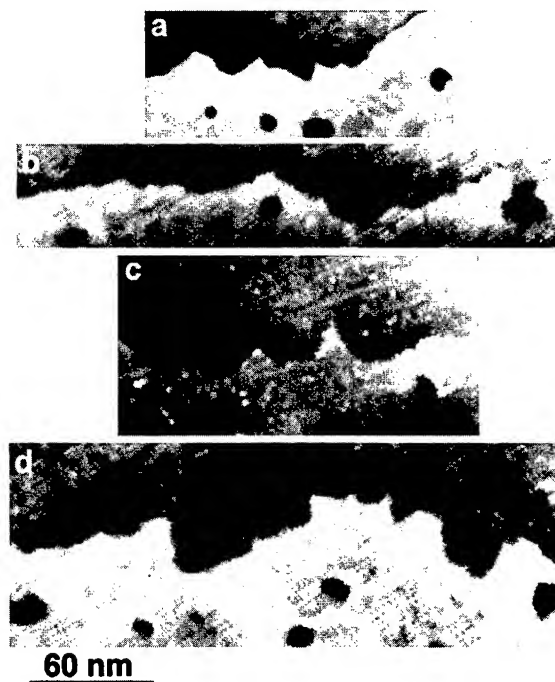


FIG. 3. Four STM images of steps, reflecting the asymmetry of voids distribution in the upper and the lower terrace.

of voids. Then a set of annealings to 600 °C was performed. STM imaging was performed after annealing for $t=0, 1, 5, 15, 30$, and 60 min. Following each STM measurement, the sample was annealed further. Two things are prominently visible in the observed images. The first is that the voids move and the second is that they attract each other and coagulate. Figure 4 displays 12 voids which are just before coagulation (a, d, and l) or immediately after coagulation (c, e, and f). In addition, we observe several voids which are clearly a result of coagulation of several small voids together (i, j, and k). These and other multiple voids seen in the images can be formed only as a result of a coagulation of several moving voids which are attracted to each other. The order of the voids (a, b, ...) increases with annealing times. The first three voids are observed after annealing for 1 min, and the last four voids after annealing for 1 h. The increase in size, associated with the coarsening process, is also seen in these multiple voids.

The void radii distributions are shown in Fig. 5. Each distribution contains information about 200 voids. The void radii for the distribution was estimated in the following way: The contrast of the image was enhanced such that the void appears black, and the upper layer appears white. Counting the total number of black pixels gives the area. From this, the average void diameter was estimated, assuming a circular shape. As was mentioned previously, voids close to the steps were not taken into account. Despite the fluctuations, it is possible to see that the average void diameter increases with time, the width of the distribution is increased also, and that the skewness of the distribution is exactly opposite to that of the ideal LSW distribution (namely, it is negative). There are

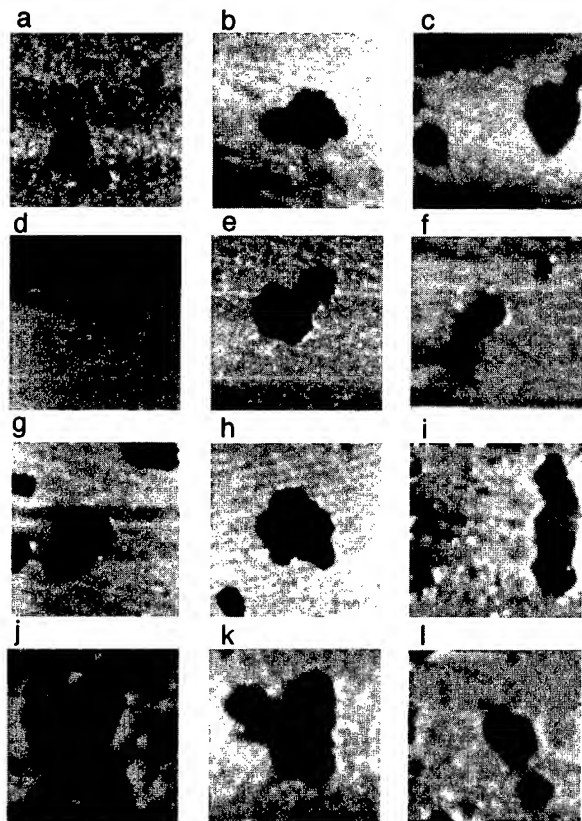


FIG. 4. Twelve STM images of voids in a Si(111) 7×7 surface. Each image is 80×80 nm².

more voids smaller than the mean than larger than the mean. This is due to the coagulation phenomenon. Two large voids may coagulate and create an even larger one. This is a mechanism used to reduce the number of voids larger than the mean, which does not exist in the ideal case. This is also

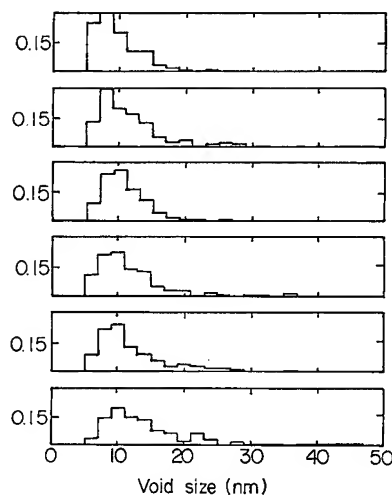


FIG. 5. Time evolution of the void size distribution in a 0.8 BL silicon layer at annealing to 600 °C. The annealing times were 0, 1, 5, 15, 30, and 60 min for the distribution functions from the top to the bottom, respectively.

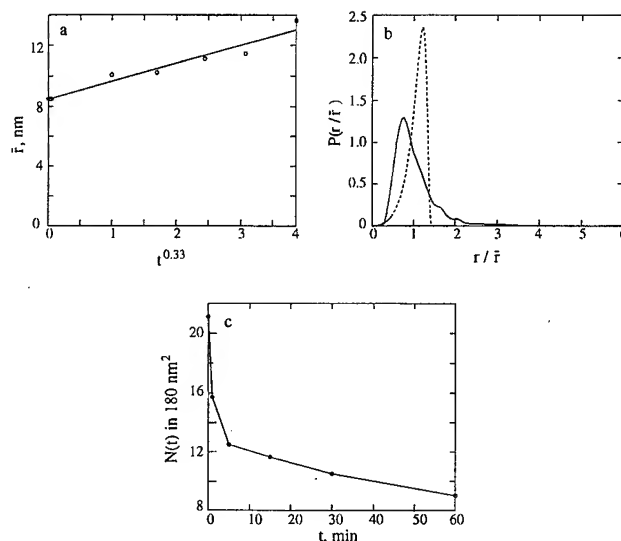


FIG. 6. (a) The increase in the average void radius taken from the distribution functions in Fig. 5 as a function of time fitted to $t^{0.33}$. (b) The sum of the distribution functions in Fig. 5 after scaling each distribution according to its average value (solid line). For comparison, the theoretical LSW distribution function (taken from Ref. 27) is also shown (dashed line). (c) The number of density of voids as a function of time.

the source of the huge voids found in the images, as can be seen in the tail of the later distributions.

To provide a somewhat clearer picture, we compared the observed results with the power law known to govern the coarsening process. In Fig. 6(a), the average void size is plotted against $t^{0.33}$. Though the data is consistent with a power law behavior of $t^{0.33}$, the dynamical range is too small, and the accuracy of the determination of the average radius is not enough to give an accurate estimate of the power law ($\alpha=0.3\pm0.1$). The small dynamic range and the fluctuations in the determination of the average void size are insufficient even to rule out a $t^{1/2}$ growth law expected for a ripening process where attachment and a detachment from the void edges is the rate limiting process. However assuming a diffusion controlled process, we can get better information on the distribution function. In order to do so, it is necessary to add the information observed from the distribution function shown in Fig. 5. The theory predicts that the distribution function obeys a scaling law. Namely that if the average radius grows with some power law $\bar{r}(t)=a\times t^\alpha$, then, the value of the distribution function at an arbitrary r is $P(r,t_2)=P(a\times t_2^\alpha)$. Thus, the distribution at later times, may be observed from an earlier distribution function just by stretching the x axis by an amount determined by the power law.

By scaling the six distribution functions in Fig. 5 with respect to their average, it is possible to observe them in one scale, and to add them together. The summed scaled normalized distribution function is shown in Fig. 6(b).

The asymmetric characteristics of the distribution with the long tail discussed earlier, is clearly distinguishable. It should be emphasized that here the small dynamic range is an advantage. Changing the scaling exponent does not sig-

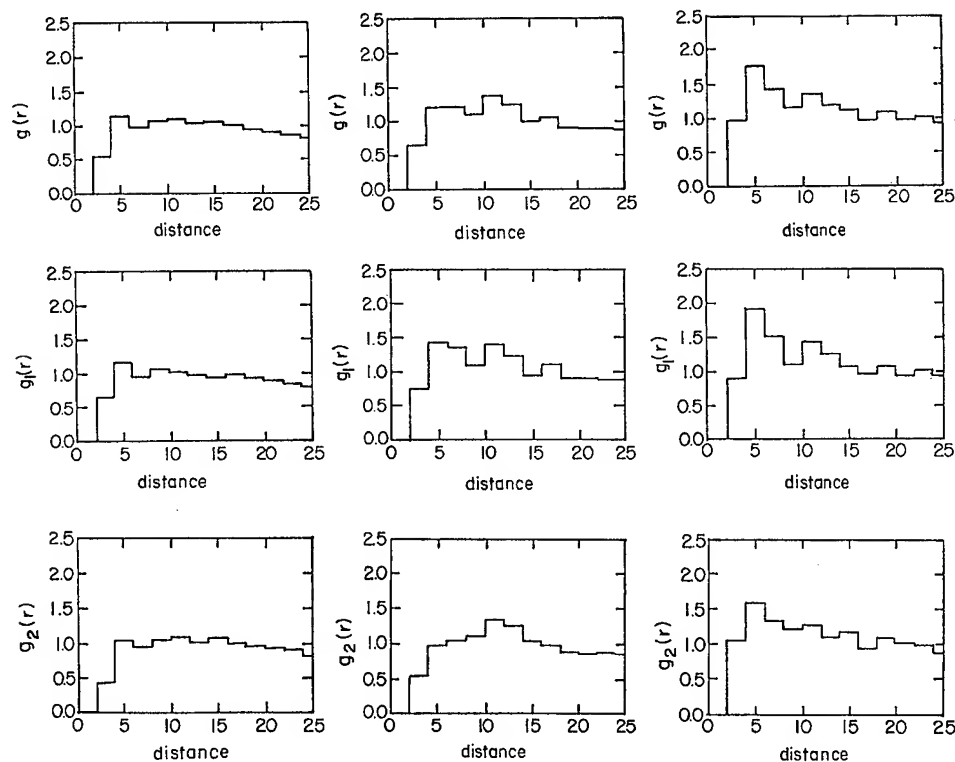


FIG. 7. (a) The pair correlation functions $g(r)$, $g_1(r)$, and $g_2(r)$ of the centers of the voids before the 600 °C annealing (left column), after annealing to 600 °C for 5 min (middle column), and for 30 min (right column). Each function represents correlations between ≈ 200 voids. The distances are given in units of r/\bar{r} .

nificantly change the width of the distribution or its shape. The sum of relative errors [the difference between the real distribution functions (Fig. 5) scaled to the average, and the distribution function shown in Fig. 6(b)] does not change much when α is changed. We conclude, therefore, that this function is a reasonable approximation for the distribution function of the radii of the voids in our experiment. However, one must be aware of the fact that the distribution functions might not be in their steady state scaling forms, as discussed later.

The increase of the average void radii, is approximately 50% during the experiments. Therefore, the average area of the void should increase by a factor of $(1.5)^2 = 2.25$. In order to keep the total area of the void constant, the number of voids (per unit area) must be reduced by the same factor. Figure 6(c) shows, that this is indeed the case, and the average number of voids (per 180 nm²-the area of most of the images) was reduced from ~ 21 to ~ 9 . This confirms a priory, that our ignorance of voids close to the step was sufficient. Otherwise the number of voids per area would have been decreased by *more* than 2.25. The observations reported here, are due to the coarsening process of the voids, alone.

In principle, coagulation may occur, if voids in a non-negligible fraction diffuse independently without interaction (random walk),^{13,14} just because of collisions. However, due to the large number of coagulated and nonspherical voids in the images, we got the impression that some attraction forces

exist between the voids. To check whether this impression is correct, we have calculated the pair correlation function of the voids:

$$g(r) = \frac{1}{2\pi r dr \rho_0} \langle n(r) \rangle,$$

where ρ_0 is the average number density of voids and $\langle n(r) \rangle$ is the average of the number of pairs of voids, to be found within a distance between r and $r + dr$. The results are shown in Fig. 7. The first, second, and third functions are related to annealing periods of 0, 5, and 30 min, respectively. In order to eliminate the effect of the increase of the average void radius on the pair correlation functions, all distances are taken relative to the average void radius. The first pair correlation function and to a lesser extent the second one, reveals no spatial correlation between the voids (and they are randomly distributed in the surface, like a gas). The third function reveals a correlation which was developed between the voids. The attraction forces between the voids caused an increase in the number of voids which are close to each other. It should be recalled, that close voids may coagulate (Fig. 4) The coagulated voids were counted as one void. The pair correlation function in Fig. 7 represents only those pairs that did not coagulate yet. To conclude, these functions prove that indeed, significant attraction forces exist between the voids and the voids, in response, migrate towards each other. The correlations do not reach a steady state!

Another piece of valuable information about the coarsening is the directional correlations which exist between the voids. This can inform us about the relative weight of the different diffusion mechanisms which might exist between the voids. On one hand, the diffusion constant is not isotropic on the Si(111)7×7 surface. However, the cross section for isotropic diffusion is larger, since all directions are available. On the other hand, the diffusion along the corner holes (or the grain boundaries) is faster, but the number of available channels for diffusion is far smaller.

To calculate the directional correlations, we have calculated the function

$$g(r, \theta) = \frac{1}{2\pi r dr d\theta \rho_0} \langle n(r, \theta) \rangle,$$

where $g(r, \theta)$ is the probability of finding two voids in between r and $r+dr$ and θ and $\theta+d\theta$. In order to get the clearest picture from $g(r, \theta)$, we applied the following integration: $g_1(r) = \int g(r, \theta) \cos^2 3(\theta - \theta_1)$, where θ_1 is the direction of the 7×7 unit vectors. This gives enhanced weighing to pairs of voids whose centers are connected by a vector in the direction of the corner holes in the 7×7 surface. In addition, we calculated $g_2(r) = \int g(r, \theta) \cos^2 3(\theta - \theta_2)$, where $\theta_2 = \theta_1 + 30^\circ$ is in between the directions of the 7×7 superlattice. g_1 and g_2 are shown in the second and third rows in Fig. 7, respectively.

Several things are distinguishable by comparing g_1 and g_2 . Some correlations do exist already after annealing for 5 min. However, hardly any difference is distinguishable between g_1 and g_2 . Namely at this stage, the diffusion is still isotropic. This makes sense, due to the fact that the smaller fraction of the surface is in the 7×7 reconstruction. In later stages of the process, as the surface is converted into the 7×7 structure, the anisotropy becomes clearer, and after 30 min, g_1 shows a much larger correlation than g_2 . This indicates that indeed, in this stage, the coarsening process occurs partly along the directions of the corner holes, either on the surface or in grain boundaries.

Although a $t^{1/2}$ ripening process could not be ruled out by our measurements of the average void size, it is more probable due to the developing spatial correlations, and their anisotropy, that the ripening process is controlled by diffusion.

IV. THEORETICAL CONSIDERATIONS

It was suggested²¹ that one characteristic of the diffusion dominated theory of movement, is attraction (or repulsion) between neighboring domains. It was shown that the two 3D domains attract each other, in the physical scenario. In two dimensions, the situation is a little more complicated, since unlike the 3D case, one cannot adjust arbitrary boundary conditions in infinity. However, one can clearly observe that in two dimensions, the interaction is much stronger. To do this, we now present a corresponding calculation in two dimensions.

We calculate the interaction between two circular domains with radii of r_1 and r_2 and a distance of $2d$ between

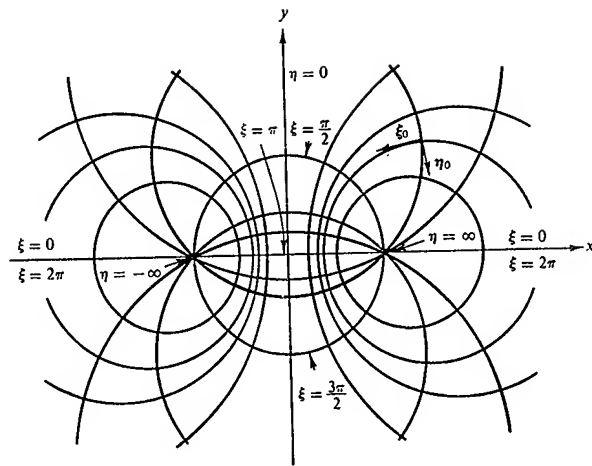


FIG. 8. The bipolar coordinates system used for the description of the 2D coarsening problem.

the centers. To solve the problem, one has to solve the static diffusion problem for the scalar diffusion field ϕ with the boundary conditions:

$$\phi_{r_1} = \frac{\gamma}{r_1}, \quad \phi_{r_2} = \frac{\gamma}{r_2},$$

where γ is the surface tension. One can calculate the velocities on the interface: $\mathbf{v}_n = \mathbf{n} \nabla \phi$, where \mathbf{n} is the normal unit vector to the interface. To know the total attraction, one has to integrate this velocity in the domain boundary.

It is convenient to use a bipolar coordinate system²⁶ where

$$x = \frac{a \sinh(\eta)}{\cosh(\eta) - \cos(\xi)} \quad \text{and} \quad y = \frac{a \cosh(\xi)}{\cosh(\eta) - \cos(\xi)}.$$

In this coordinate system, $\eta_r = \text{const}$ are curves along circular trajectories with radius r (see Fig. 8). It is easy to show that

$$a = \sqrt{d^2 - r^2} \quad \text{and} \quad \exp(2\eta_r) = \frac{d + \sqrt{d^2 - r^2}}{d - \sqrt{d^2 - r^2}}.$$

The equation $\nabla^2 \phi = 0$ is separable in the bipolar coordinates. The problem becomes a problem of two lines η_{r_1} and η_{r_2} with potentials ϕ_1 and ϕ_2 . The obvious solution is

$$\phi = \phi_2 + \frac{(\phi_1 - \phi_2)}{\eta_{r_1} - \eta_{r_2}} (\eta - \eta_{r_1}).$$

We note that infinity in (x, y) maps into the point $(0, \pi)$ in the bipolar coordinate frame. The only possible solution with no flow to infinity, which is the physical condition, is our result. To resolve other possibilities, one needs to fix boundary conditions properly which we cannot do in this case.

The gradient on the η_{r_1} plane becomes

$$\nabla_{r_1} \phi = \frac{a}{\cosh \eta_{r_1} - \cos \xi} \frac{\phi_1 - \phi_2}{\eta_{r_1} - \eta_{r_2}}.$$

To find the total velocity, one has to integrate $\nabla\phi$ over the domain boundaries. We perform the integration only for $d \gg r$. In this case, ξ becomes the same as θ (the integral angle) and

$$v \approx \frac{a}{2\pi r_1 \eta_{r_1} - \eta_{r_2}} \int \frac{\cos \theta}{\cosh \eta_{r_1} - \cos \theta} d\theta.$$

Using the equations for a and $\exp(2\eta_r)$, we get the average velocities:

$$v_{\text{tot}}^1 \approx \frac{\gamma(1/r_1 - 1/r_2)}{2\pi[\log(r_1/d) + \log(r_2/d)]} \frac{r_1}{d}.$$

The velocities are in the same directions, i.e., the large cluster moves toward the smaller one and a smaller one away from the larger one. Thus indeed, as expected, domains move in the direction of the larger concentration of vacancies because this gradient generates a difference in the growth rates in the two sides. Though we find that the velocity of the larger one is higher, this calculation does not explain the phenomena observed in our experiment. To see why, we refer to the equation for the gradient of the field. The velocities on the edges are either all positive or all negative. This indicates that though a movement of the center of mass is possible, the cluster will never cut its original boundaries. This can be clearly observed in the two dimensional simulations of this type of process.²¹ Furthermore, since the total mass that can be transferred to r_2 from the smaller cluster r_1 is πr_1^2 . The cluster can move at most r_1 which is a very short distance, and does not seem to match our experimental observation.

To give an explanation to the strong domain migration, we first remind the reader that aside from diffusion between the clusters there is an additional process with a much lower threshold which is the release of adatoms from a step. This process is more dominant in larger voids, since the interface is mostly flat. This process is quite fast without surface tension as observed before at the step.

Though we do not know what the main transport mechanism between the voids, adatoms, or vacancies is, we assume for the sake of our argument that it is the second possibility.

Consider a large void and suppose that a vacancy was to reach the void. The result will be a creation of a missing atom on the void boundary. The atoms near this point have a much higher probability of emission into the void. So, a single vacancy entrance into the void will destabilize a whole region. The adatom emitted from this region will be adsorbed on less strained parts of the void boundary. This is a mechanism that enhances very strongly the response to a single vacancy entrance.

If this mechanism is true, it explains the large movement and the strong distortions observed during the process. A

similar type of arguments might be also made in relation to adatom jump mechanism.

In summary, we observed a coarsening process of voids. We observe that unlike simple diffusion dominated processes, usually described by LSW type of models, a strong domain mobility and attraction is observed in the experiment.

We reason that a possible mechanism for the strong void mobility and attraction is boundary destabilization by vacancy migration.

This picture is consistent with our experimental observation. It is an interesting example for the importance of microscopic details in the global dynamics of coarsening.

ACKNOWLEDGMENTS

This work was supported by the Minerva Foundation, Munich, Germany, and the Basic Research Foundation administered by the Israeli Academy of Sciences and Humanities. One of the authors (Y. M.) is an incumbent of the Lilian and George Lytle Career Development Chair.

¹W. Ostwald, Z. Phys. Chem. (Munich) **34**, 495 (1900).

²I. S. Servi and D. Turnbull, Acta Metall. **14**, 161 (1966).

³P. W. Voorhees and R. J. Schaefer, Acta Metall. **35**, 327 (1987).

⁴M. Bahiana and Y. Oono, Phys. Rev. A **41**, 6763 (1990).

⁵P. Bassereau, D. Brodbeck, T. P. Russel, H. R. Brown, and K. R. Shull, Phys. Rev. Lett. **71**, 1716 (1993).

⁶X. Zheng and B. Bigot, J. Phys. (France) II **4**, 743 (1994).

⁷W. Y. Shih, W.-H. Shih, and I. A. Aksay, J. Mater. Res. **10**, 1000 (1995).

⁸O. Krichevski and J. Stavans, Phys. Rev. Lett. **70**, 1473 (1993).

⁹M. Seul, N. Y. Morgan, and C. Sire, Phys. Rev. Lett. **73**, 2284 (1994).

¹⁰G.-C. Wang and T.-M. Lu, Phys. Rev. Lett. **50**, 2014 (1983).

¹¹A. Stabel, R. Heinz, F. C. De Schryver, and J. P. Rabe, J. Phys. Chem. **99**, 505 (1995).

¹²G. Coulon, B. Collin, D. Chatenay, and Y. Gallot, J. Phys. (France) II **3**, 697 (1993).

¹³J.-M. Wen, S.-L. Chang, J. W. Burnett, J. W. Evans, and P. A. Thiel, Phys. Rev. Lett. **73**, 2591 (1994).

¹⁴K. Morgenstern, G. R. Rosenfeld, B. Poelsema, and G. Comsa, Phys. Rev. Lett. **74**, 2058 (1995).

¹⁵I. M. Lifshitz and V. V. Slyozov, J. Phys. Chem. Solids **19**, 35 (1961).

¹⁶C. Wagner, Z. Elektrochem. **65**, 581 (1961).

¹⁷J. A. Marqusee, J. Chem. Phys. **81**, 976 (1984).

¹⁸J. A. Marqusee and J. Ross, J. Phys. Chem. **80**, 536 (1984).

¹⁹M. Marder, Phys. Rev. Lett. **55**, 2953 (1985).

²⁰Q. Zheng and J. D. Gunton, Phys. Rev. A **39**, 4848 (1989).

²¹P. W. Voorhees, G. B. McFadden, R. F. Boisvert, and D. I. Meiron, Acta Metall. **36**, 207 (1988).

²²W. Pflugel and U. M. Titulaer, Physica A **214**, 52 (1995).

²³M. Zinke-Allmang, L. C. Feldman, and M. H. Grabow, Surf. Sci. Rep. **16**, 377 (1992).

²⁴E. Ter-Ovanesyan, Y. Manassen, and D. Shachal, Phys. Rev. B **50**, 8020 (1994).

²⁵E. Ter-Ovanesyan, Y. Manassen, and D. Shachal, Isr. J. Chem. **36**, 45 (1996).

²⁶G. Arfken, *Mathematical Methods for Physicists* (Academic, New York, 1966).

²⁷T. M. Rogers and R. C. Desai, Phys. Rev. B **39**, 11 956 (1989).

Liquid-phase adsorption process of β -picoline on stilbite (010) observed by atomic force microscopy

Masaharu Komiyama^{a)} and Minming Gu^{b)}

Department of Chemistry, Yamanashi University, Takeda, Kofu 400, Japan

(Received 8 September 1996; accepted 1 April 1997)

Liquid-phase adsorption process of β -picoline molecules on a (010) surface of a natural zeolite stilbite was observed by atomic force microscopy (AFM). Initially, the surface was covered with excess multilayer β -picoline molecules. Removal of this excess layer by AFM scanning exposed a possible random-adsorbed monolayer phase. A well-ordered two-dimensional quasihexagonal phase developed slowly, with an estimated rate of ~ 6 nm/min. The 2D unit cell dimension of the latter ordered phase was 0.55 ± 0.04 nm. The ordered phase is incommensurate with the substrate surface, although one of the 2D lattice unit vectors of the adsorbed phase makes an angle of $\sim 50^\circ$ and $\sim 100^\circ$ with respect to the a and c axes of the substrate (010) plane, respectively. From the position of the methyl group in the β -picoline molecule it is estimated that the molecules are adsorbed on the surface with their C–N bond parallel to the surface. © 1997 American Vacuum Society. [S0734-211X(97)02904-1]

I. INTRODUCTION

Molecular adsorption on solid surfaces has been the target of extensive studies in the field of surface science and catalysis. The packing structures and molecular orientations of the adsorbed molecules have been established in many adsorbate-substrate systems. The vast majority of this work, however, has been limited to metal or semiconductor substrates under high or ultrahigh vacuum conditions, mainly due to the requirements by the techniques employed to obtain such information. Furthermore, the adsorption processes under these conditions are quite fast, and the dynamics of the adsorption process are not commonly observable.

The advent of scanning probe microscopy (SPM) removed some of these requirements: scanning tunneling microscopy (STM) extended the examination of adsorption from high vacuum to higher pressures and even under liquid phase conditions,^{1–4} and atomic force microscopy (AFM) enabled the examination of adsorption onto insulating substrate systems.^{5–8}

As part of our investigations on the liquid-phase adsorption of pyridine base molecule on zeolite surfaces, we report here the examination of the adsorption of β -picoline on a cleaved stilbite (010) surface by AFM. In our previous work,^{7,8} we examined pyridine adsorption on two natural zeolite surfaces, heulandite (010) and stilbite (010). In the work the formation of well-ordered hexagonal or quasihexagonal pyridine adlayers on these surfaces were observed. While we were able to determine the array structures and to estimate the molecular orientations within the pyridine adlayers, a relatively fast rate of the ordered layer formation prevented us from examining the detailed process of the development of the adlayers. We were only able to note that it takes some time for the ordered layer to develop, and before

that we observed some intermediate phase between the clean substrate surface and the ordered adlayer phase.^{7,8}

The present work found that the rate of the ordered β -picoline adlayer formation on stilbite (010) is very slow, and a possible intermediate adsorption phase exists on the surface, likely to be a randomly adsorbed one, before the emergence of a highly ordered one. The development rate of the latter phase was determined, as well as its array structure and registry with respect to the substrate surface structure.

II. EXPERIMENT

Stilbite, from Bear Brook, Nova Scotia, Canada, was used as a zeolite substrate. Its nominal composition is $\text{Na}_4\text{Ca}_8[\text{Al}_{20}\text{Si}_{52}\text{O}_{144}] \cdot 56\text{H}_2\text{O}$, with the unit cell lattice parameters $a = 1.36$ nm, $b = 1.82$ nm, $c = 1.13$ nm, and $\beta = 128^\circ$. Micropores consisting of ten-membered rings run along the [100] direction, viz., parallel to the (010) (i.e., ac) plane that constitutes the fault plane which is exposed by cleaving. Along the b axis two of such fault planes exist, giving a step height half the b unit length when cleaved.

The sample was cleaved with a razor blade along its (010) [or (020)] fault plane under ambient conditions and placed in a sealed AFM liquid cell which is then filled with deionized and membrane-filtered water. The AFM examinations were performed using a Nanoscope II contact-mode AFM (Digital Instruments) in the repulsive force range with a tip load of ~ 5 nN unless noted otherwise. Type NP cantilevers (Digital Instruments) were employed, which have Si_3N_4 tips and force constant equal to 0.38 kN/m. All the recorded images are presented without filtering.

III. RESULTS AND DISCUSSION

The cleaved stilbite surface gives extremely flat terraces extending more than a few hundred nanometers,⁷ each separated by an ~ 1 nm step, a height corresponding to the stilbite (020) lattice spacing that is 0.91 nm (cf. Fig. 2). Figure

^{a)}Electronic mail: komiyama@chem.ksb.yamanashi.ac.jp

^{b)}Present address: Physics Department, Centre for the Physics of Materials, McGill University, Montreal, Quebec H3A 2T8, Canada.

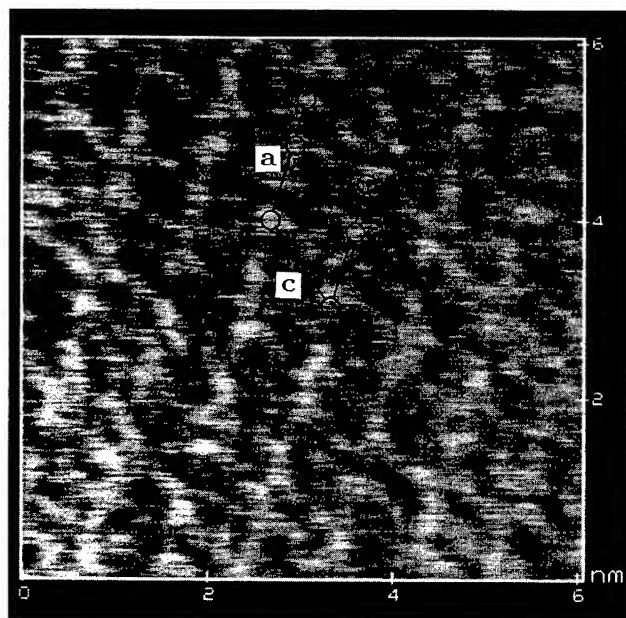


FIG. 1. A narrow-area AFM image of a cleaved stilbite (010) surface obtained under water. The grey scale is 0.5 nm full scale. A surface unit cell expected for the bulk-terminated (010) face is inset. Open circles in the unit cell indicate the positions of top-layer oxygen atoms.

1 shows an 8 nm \times 8 nm, narrow-area scan of the cleaved stilbite (010) surface obtained under water. Superposed on the image is the surface unit cell expected for a bulk-terminated (010) stilbite surface, with open circles indicating the topmost oxygen atoms. On the a axis two oxygen atoms per unit cell are present with unequal spacing, 0.83 nm for the longer pitch and 0.53 nm for the shorter. While the position of each oxygen atom is not clear in the observed AFM

image, observed surface translational symmetry (particularly in the c direction) corresponds quite well with the unit cell of the bulk terminated (010) surface.

After images under water were taken, the liquid content in the cell is substituted with 1 vol % aqueous β -picoline solution. Figure 2 shows 970 nm \times 970 nm, wide-area images of the surface observed under β -picoline solution. In Fig. 2(a) fuzzy "clouds" are observed on the surface. At the center of the image there exists a region devoid of the clouds. This region corresponds to the area of preceding narrow-area scanning. In fact two more scans on this area removed the clouds entirely as may be found in Fig. 2(b), exposing a terraced surface that is macroscopically no different from images obtained under pure water. The substance of the clouds is attributed to the large and excess quantity of β -picoline molecules, precipitated on the surface.

Detailed AFM examination of this cloud-removed surface gives images somewhat different from that of the clean surface obtained under pure water. An example is shown in Fig. 3, which was taken 1 h after the image shown in Fig. 2 was taken. This AFM image still holds the translational symmetry of the stilbite (010) surface as indicated by the superposed (010) unit cell, but shows other structures that are not attributable to atomic arrangements expected for the clean surface. Although it is difficult to discriminate artifacts due to tip apex changes from actual surface structure changes, we attribute this as a new surface structure, emerged by the randomly chemisorbed β -picoline molecules. In the image, there are structures that might be attributable to individual picoline rings, but their images are not clear enough to identify them as individual adsorbed molecules. The height variation in AFM images expected for a flatly laying picoline molecule is very small. For a similar molecule, pyridine, adsorbed on heulandite we observed with AFM less than a

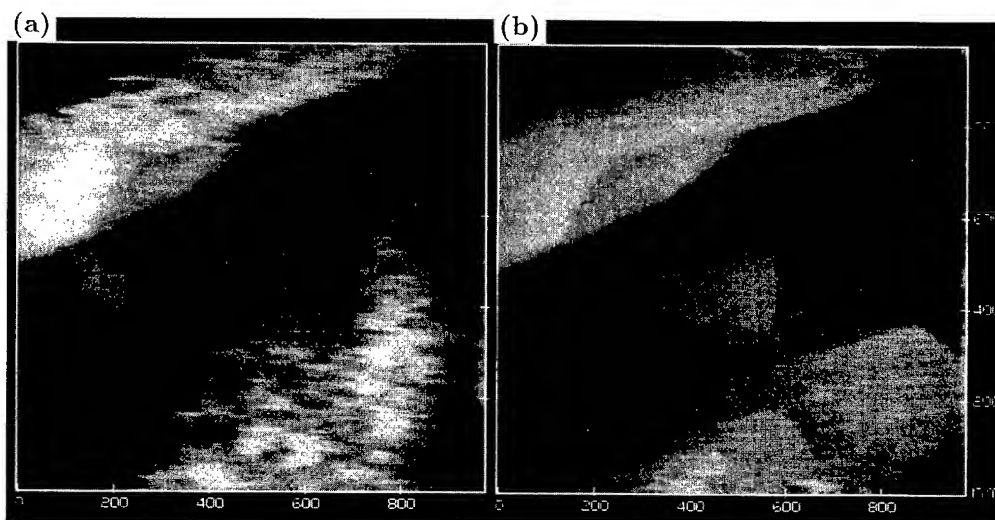


FIG. 2. Wide-area AFM images of a cleaved stilbite (010) [or (020)] surface obtained under β -picoline solution. First scan gives a "cloud"-covered surface image (a), whereas the cloud is swept away with two more scans on the surface, leaving a smooth surface (b). The exposed surface is extremely flat, with terraces over a few hundred nanometers wide, separated by steps whose height of ~ 1 nm corresponds to half of the b axis unit cell length of 1.82 nm. The grey scale is 10 nm full scale.

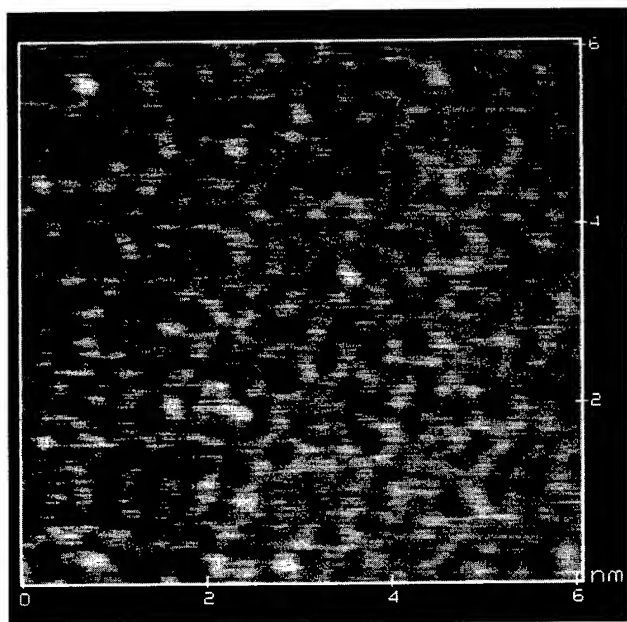


FIG. 3. A narrow-area AFM image of a stilbite (010) observed under aqueous β -picoline solution taken 1 h after the images shown in Fig. 2 were taken. The image shows different surface structure from that taken under pure water (Fig. 1) but still shows translational symmetry of the stilbite (010) surface as may be found by the inset of the (010) unit cell. The grey scale is 0.5 nm full scale.

0.1 nm height variation when the molecules are overlapping each other.⁸ In STM, surface corrugations are often dependent on bias voltages, but Ohtani *et al.*⁹ report benzene corrugation between 0.05 and 0.2 nm at vacuum-Rh (111) interface, and Weiss and Eigler¹⁰ between 0.06 and 0.09 at a vacuum-Pt {111} interface. Yau, Kim, and Itaya report a STM value of 0.07 nm for benzene molecules adsorbed on liquid-Pt(111) or Rh(111) interfaces.⁴ Thus it was not possible to identify the present picoline molecules by the height

analysis of the image shown in Fig. 3. Very small height modulation on the image, on the other hand, and the fact that we can still see the translational symmetry of the substrate surface in the image, indicate that the molecules are adsorbed as a monolayer and no multilayer adsorption exist under the present AFM observation conditions.

Three hours after the image in Fig. 3 was taken, an image with a different structure was obtained as shown in Fig. 4. At the upper-left corner of Fig. 4(a) the structure identical to the one observed in Fig. 3 may be seen, as indicated by the inset of the (010) unit cell, whereas at the lower-right quarter a different structure is being developed. As may be found in the next scan taken after 40 s, shown in Fig. 4(b), this new phase has an ordered quasi-hexagonal structure. This (quasi-) hexagonal structure is also characteristic of pyridine molecules adsorbed on zeolite(010) surfaces.^{7,8}

Comparison of the sequential images shown in Figs. 4(a) and 4(b) indicates that the ordered phase expanded itself within the 40 s time interval. In Fig. 4(a) the ordered phase is confined to the lower-right quarter, whereas in Fig. 4(b) the ordered area has expanded itself to leave only the upper-left quarter for the disordered phase. Further scanning found that the imaged area is completely filled with this ordered phase. From these images it is possible to estimate the development rate of the ordered phase to be ~ 6 nm/min. This rate of ordered layer development is about two orders of magnitude slower than that of porphyrin on iodine-modified Ag(111) observed by STM.²

The above observation may indicate that the ordering process is a transition from the disordered to the ordered phase, which may be initiated at some nucleation sites. The identity and the nature of these nucleation sites, however, are not apparent at this moment. Further examination toward this point may be necessary.

A close examination of the AFM images presented in Fig. 4 also allows us to determine the registry of the ordered

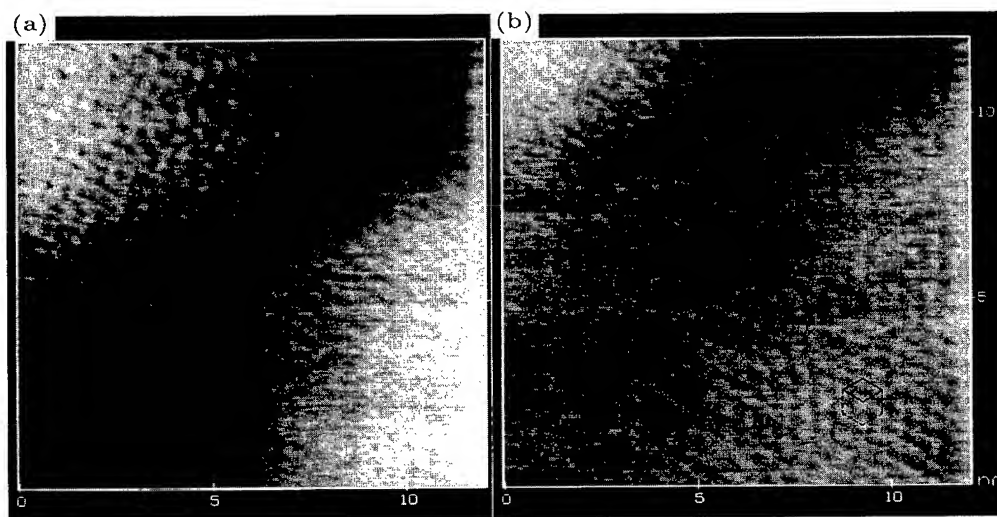


FIG. 4. Sequential AFM images of a β -picoline-adsorbed stilbite (010) surface taken 3 h after image shown in Fig. 3 was taken, showing the development of the ordered phase. Time interval between the images (a) and (b) was 40 s. In the randomly adsorbed area the (010) unit cell is overlaid, and in the ordered area a unit cell for the hexagonal array is inset. The grey scale is 0.4 nm full scale.

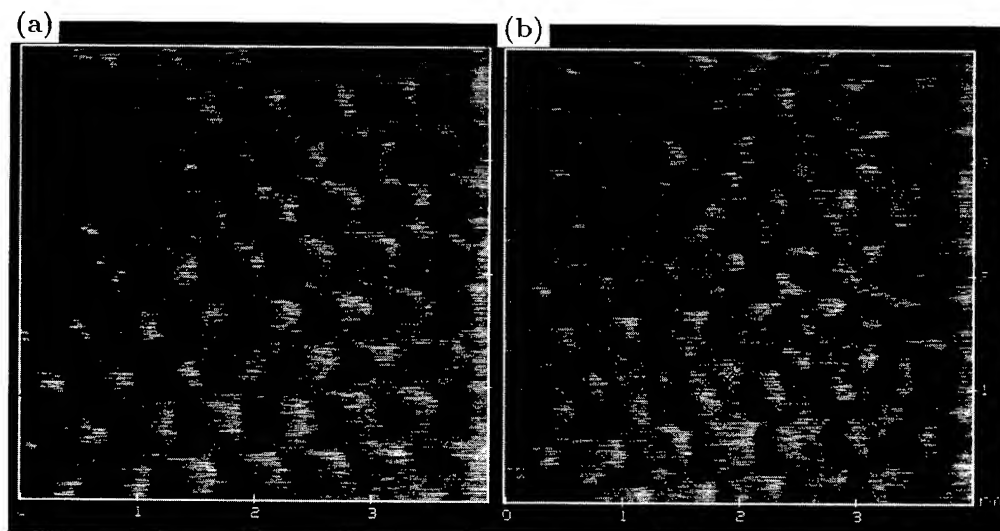


FIG. 5. Narrow-area images of the ordered phase. Image (a) was taken with ~ 5 nN tip load, and (b) with ~ 50 nN. A ball-and-stick model of β -picoline was overlaid on one ring image. The grey scale is 0.50 nm full scale.

adlayer with respect to the substrate surface, which was not possible for pyridine adsorption.^{7,8} The a and c axes of the substrate, as determined following the case of Fig. 3, and the rhombic unit cell of the ordered adlayer, are inset in Fig. 4. Comparison of the two unit cells indicates that one of the adlayer lattice vectors makes angles of $\sim 50^\circ$ and $\sim 100^\circ$ with respect to the a and c axes of the stilbite(010) plane, respectively. It is also noted that height analysis in Fig. 4 indicates that the picoline adlayer consists most likely of a single molecular layer. There exist no appreciable height differences (within 0.1 nm) between the ordered and the random regions.

As in the case of pyridine on stilbite(010) and on heulandite(010), the ordered structure of the β -picoline adlayer does not have strong correlation with the periodicity of the stilbite surface unit cell. For pyridine adlayers we concluded that the adlayer is incommensurate with both the stilbite(010) and heulandite(010) surface lattices,^{7,8} and this appears to be true for the case of β -picoline on stilbite(010). Thus, for β -picoline we also assume molecular self-assembly as the driving force for the formation of the ordered adlayer. The ordered layer, which may be considered as a 2D solid,⁷ then interacts with the zeolite surface collectively. While there exist a number of reports on the structure of pyridine adlayers on single-crystal metal surfaces (see literature cited in Ref. 8), no reports on the structure of picoline adlayer on metal surfaces, either under vacuum or aqueous conditions, are known to the authors.

When scanning was confined to this ordered region, images shown in Fig. 5 are obtained. Figure 5(a) was taken with the tip load of ~ 5 nN, and Fig. 5(b) with ~ 50 nN. The surface appears to be covered with a well-ordered array of rings similar to the pyridine adsorption on zeolite surfaces.^{7,8} A ball-and-stick model of a β -picoline molecule is overlaid on one ring structure. Here the picoline plane is placed parallel to the surface, although tilting it to some degree, 47° to the surface for instance (see later), does not affect the fol-

lowing observations. The size of the observed ring appears to be well-matched with the molecular model of β -picoline, supporting the notion that the observed image is that of the adsorbed picoline molecules. As may be seen from the images, the array structure shows almost hexagonal symmetry. From the AFM images presented and not presented in this article, the unit cell dimension was determined to be 0.55 ± 0.04 nm and the angle between the two lattice vectors to be $60^\circ \pm 5^\circ$.

The molecules found in Fig. 5 appear to take a tilted position similar to the case of pyridine molecules.^{7,8} This is apparent in Fig. 5(a), where the molecular images show only half of the rings. This is interpreted as due to the fact that the molecule is taking a tilted position with respect to the surface, and only the upper part of the ring is probed by the AFM tip. If we consider, to the first approximation as in the case of pyridine on heulandite(010),⁸ that the density of the adsorbed phase equals that of a β -picoline solid,¹¹ the tilt angle is calculated to be 47° to the surface. This value is very close to the similarly estimated tilt angle of pyridine on heulandite(010) which is 43° .⁸

It is also noted in Fig. 5 that imaging of the molecular inner structure seems to be dependent on the tip load. With a low (~ 5 nN) tip load the molecules appear as simple rings, with their lower halves not being visible [Fig. 5(a)]. With a high (~ 50 nN) tip load the entire ring for each molecule is apparent [Fig. 5(b)]. Furthermore, in this high load image there exists a protrusion attached to each ring image. We consider that this part represents the methyl group of each picoline molecule, as indicated by the skeleton model inset. From this image we were able to estimate the configuration of the picoline molecules with respect to the substrate surface. Judging from the position of the methyl group, two configurations are possible: one with a C–N bond parallel to the surface, and the other with the bond between No. 5 and 6 carbon atoms of the aromatic ring being parallel to the surface.

- ¹M. Kunitake, N. Batina, and K. Itaya, *Langmuir* **11**, 2337 (1995).
²K. Ogaki, N. Batina, M. Kunitake, and K. Itaya, *J. Phys. Chem.* **100**, 7185 (1996).
³N. Batina, M. Kunitake, and K. Itaya, *J. Electroanal. Chem.* **405**, 245 (1996).
⁴S.-L. Yau, Y.-G. Kim, and K. Itaya, *J. Am. Chem. Soc.* **118**, 7795 (1996).
⁵A. L. Weisenhorn, J. E. MacDougall, S. A. C. Gould, S. D. Cox, W. S. Wise, J. Massie, P. Maivald, V. B. Elings, G. D. Stucky, and P. K. Hansma, *Science* **247**, 1330 (1990).
⁶H. Cai, A. C. Hillier, K. R. Franklin, C. C. Nunn, and M. D. Ward, *Science* **266**, 1551 (1994).
⁷M. Komiya and M. Gu, *Jpn. J. Appl. Phys.* **1** **35**, 3775 (1996).
⁸M. Komiya, T. Shimaguchi, T. Koyama, and M. Gu, *J. Phys. Chem.* **100**, 15198 (1996).
⁹H. Ohtani, R. J. Wilson, S. Chiang, and C. M. Mate, *Phys. Rev. Lett.* **60**, 2398 (1988).
¹⁰P. S. Weiss and D. M. Eigler, *Phys. Rev. Lett.* **71**, 3139 (1993).
¹¹S. G. Biswas, *Indian J. Phys.* **35**, 261 (1961).

Magnetic nanostructures studied by scanning probe microscopy and spectroscopy

R. Wiesendanger, M. Bode, M. Kleiber, M. Löhndorf, R. Pascal, and A. Wadas
*Institute of Applied Physics and Center for Microstructure Research, University of Hamburg,
Jungiusstrasse 11, D-20355 Hamburg, Germany*

D. Weiss

*Max-Planck-Institute for Solid State Research, Stuttgart, Germany and Institute of Experimental
and Applied Physics, University of Regensburg, Regensburg, Germany*

(Received 8 September 1996; accepted 15 February 1997)

Scanning tunneling microscopy and spectroscopy (STM/STS) were applied to study the nanostructural and nanoelectronic properties of ultrathin magnetic films, from individual adatoms and clusters (zero-dimensional systems) to three-dimensional bulklike islands. Complementary investigations of the nanomagnetic structure of ultrathin films were performed by magnetic force microscopy (MFM). Quantum magnetic structures were prepared by electron-beam lithography and nanosphere lithography and analyzed by STM/STS as well as by MFM. Magnetization switching of single-domain Co dots induced by the MFM tip is demonstrated, and its potential application for quantum magnetic storage is shown. Finally, tunneling spectroscopy performed with magnetic probe tips yields new information about the spin-resolved nanoelectronic properties of magnetic nanostructures. © 1997 American Vacuum Society. [S0734-211X(97)05704-1]

I. INTRODUCTION

Magnetic nanostructures have become of considerable interest because of recent developments in magnetic sensor technology and quantum magnetic storage. Vertical magnetic nanostructures such as thin films and multilayer systems were first the focus of research and they have led to new discoveries such as giant magnetoresistance.¹ More recently, lateral magnetic nanostructures have been fabricated by several different methods, such as electron beam lithography² or nanoimprint lithography.³ Scanning probe microscopy and spectroscopy is ideally suited to address questions related to the relationship between topographic, electronic, and magnetic properties of magnetic nanostructures which are of particular interest in order to tailor their magnetic behavior for various applications. In the present work, we have applied scanning tunneling microscopy (STM) and spectroscopy (STS) to investigate structural and electronic properties of magnetic nanostructures, from individual adsorbates to three-dimensional (3D) bulklike islands. The magnetic properties of magnetic nanostructures fabricated by electron-beam lithography and nanosphere lithography were studied by magnetic force microscopy (MFM). Finally, we have addressed the relationship between electronic and magnetic properties by performing comparative tunneling spectroscopy studies with nonmagnetic and magnetic probe tips.

II. EXPERIMENT

The STM/STS measurements were performed with a modified commercial instrument⁴ in ultrahigh vacuum (UHV) with a base pressure in the low- 10^{-11} mbar range. The UHV chamber was additionally equipped with sample preparation (heating stage, thin film evaporator) and surface

analysis (low energy electron diffraction, Auger electron spectroscopy) facilities. MFM measurements were performed either under ambient conditions with a commercial instrument⁵ or under UHV conditions with a recently developed UHV-MFM⁶ based on a modified commercial instrument.⁷ For ambient MFM studies we used single crystal silicon cantilevers with integrated probe tips⁸ coated under high vacuum with a thin layer (20–30 nm) of Fe and a protective coating of Au. For UHV-MFM studies as well as for magnetic-sensitive tunneling spectroscopy studies we have coated our probe tips (single crystal silicon and electrochemically etched tungsten tips) *in situ* by an Fe overlayer. In this case, a protective coating was not necessary because the magnetically coated tips were subsequently transferred *in situ* onto the microscope stage without breaking the UHV conditions. For each set of measurements, several different probe tips were prepared to confirm that the experimental results presented below are independent of microscopic details of the tip structure.

III. STM/STS STUDIES OF ZERO-THREE DIMENSIONAL MAGNETIC NANOSTRUCTURES

To investigate the structural and electronic properties of magnetic nanostructures from individual zero-dimensional (0D) adatoms up to 3D bulklike islands we have performed a complementary STM/STS study of Fe on W(110) at different coverages. Individual Fe adsorbates were observed upon evaporation of a small amount of iron [<0.001 monolayers (ML)] onto a (15×3)-reconstructed W(110) single crystal substrate. On nonreconstructed W(110) substrates, we were never able to observe individual iron adatoms. This can be explained by a relatively low barrier for surface diffusion of Fe atoms on W(110)-(1×1) compared with the thermal en-

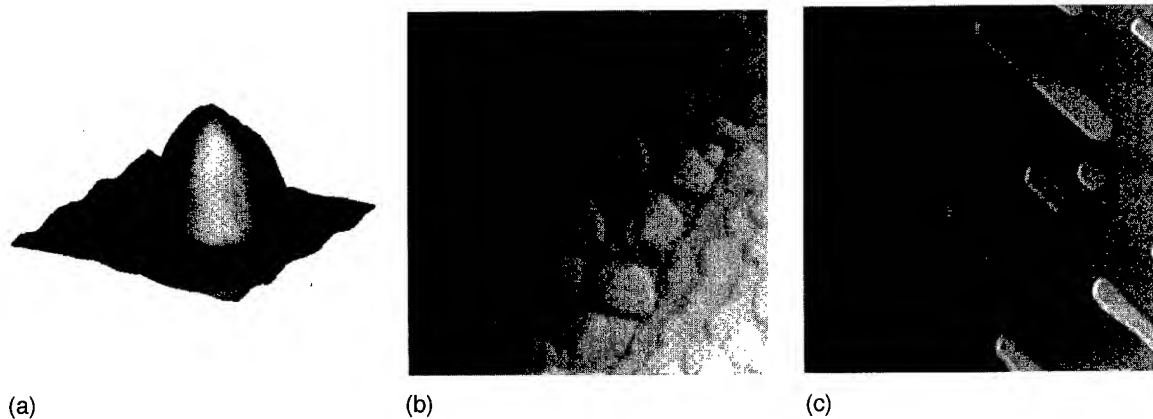


FIG. 1. (a) Single iron nanocluster on the W(110)/C-(15 \times 3) substrate at 300 K. (Apparent height: 0.23 nm, FWHM: 1.39 nm). (b) Topographic STM image (60 nm \times 60 nm) of 2D iron islands on terraces of a stepped W(110) substrate. (c) Topographic STM image (500 nm \times 500 nm) of 3D iron islands on a stepped W(110) substrate. The islands are about 40 ML high. Between them, a single iron monolayer covers the bare W(110) substrate.

ergy at 300 K. Consequently, low temperatures are required to image individual atoms on nonreconstructed metal surfaces, as demonstrated, e.g., for Fe on Pt(111) (Ref. 9) or Fe on Cu(111).¹⁰ Alternatively, surface reconstructions induced, e.g., by small amounts of carbon usually enhance the surface potential corrugation, thereby preventing the adsorbed atoms from diffusing on the substrate surface at 300 K. They can thereby be imaged by STM. The individual Fe adatoms appear with a height of 0.15 nm and a full width at half-maximum (FWHM) of 0.9 nm in constant-current scans at 300 K. This is in reasonable agreement with low temperature data⁹ that take small thermal broadening effects into account. We have also observed nanometer-scale iron adclusters as shown in Fig. 1(a). In Fig. 2(a) we present local tunneling spectra (dI/dU - U characteristics) measured above the W(110)/C-(15 \times 3) substrate (A), above an individual iron adsorbate (B), and above a nonmetallic adsorbate (C). A pronounced peak at a sample bias voltage of +0.5 V is observed in the spectrum measured above the iron adsorbate,

corresponding to an unoccupied electronic state 0.5 eV above the Fermi energy E_F . This empty-state feature appears to be characteristic of individual iron adatoms and adclusters and was discovered earlier in low temperature STM studies of Fe on Pt(111).⁹ It is interesting that the observed peak in the tunneling spectrum 0.5 eV above E_F for Fe adsorbates seems to be independent of the chosen substrate (and of the chosen tip), indicating that this peak can provide a spectroscopic fingerprint for the identification of iron adsorbates.¹¹

Quasi-two-dimensional (2D) islands of iron were prepared by evaporating a submonolayer onto a clean W(110) substrate held at room temperature. Figure 1(b) shows a topographic STM image of the 2D iron islands with a height of 1 ML. Upon positioning the probe tip above an iron island, the local tunneling spectrum (B) of Fig. 2(b) was obtained. Again, a pronounced peak at a sample bias voltage of $U = +0.2$ V can be observed, similar to that of individual iron adsorbates [Fig. 2(a)], but slightly shifted towards lower

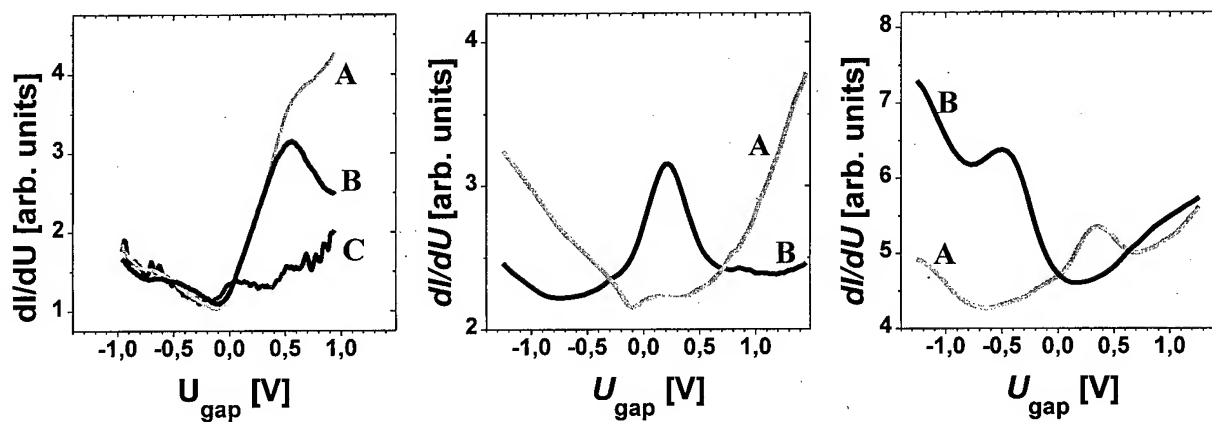


FIG. 2. (a) Local tunneling spectra (dI/dU vs U) obtained for three different positions of the probe tip: (A) above the W(110)/C-(15 \times 3) substrate, (B) above an iron adsorbate, (C) above a nonmetallic adsorbate. A pronounced empty-state peak at $U = +0.5$ V appears in spectrum (B). (b) Local tunneling spectra measured above the W(110) substrate (A) and above the 2D iron islands (B). A pronounced empty-state peak at $U = +0.2$ V appears in spectrum (B). (c) Local tunneling spectra obtained above the single iron monolayer (A) and above the 3D iron islands (B). A pronounced filled-state appears at $U = -0.5$ V in spectrum (B), whereas the empty-state peak characteristic of a single monolayer [spectrum (A) and Fig. 2(b)] is totally absent in (B).

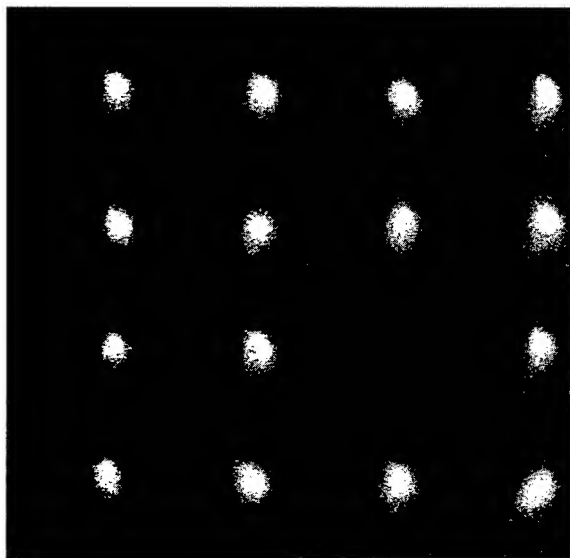


FIG. 3. MFM image ($2\ \mu\text{m} \times 2\ \mu\text{m}$) of a Co dot array on a GaAs substrate fabricated by electron-beam lithography. Applying an external field of $H_{\text{ext}} = 3000\ \text{Oe}$ led to a uniform magnetized state of the single domain particles.

energy. For comparison, the local tunneling spectrum measured above the bare W(110) substrate [curve A in Fig. 2(b)] is shown. The minimum in the local dI/dU - U characteristic around zero bias for the W(110) substrate is due to a low density of states close to E_F and is well known for this particular surface. The pronounced difference between the spectroscopic curves for the iron islands and the W(110) substrate provides ideal conditions for element-specific imaging by STS.¹²

Upon evaporation of nominally 8 ML of iron on W(110) at room temperature, a relatively rough surface topography of the iron thin film was observed. By annealing the film at about 800 K, the iron piles up and forms 3D islands with atomically well defined (110) facets on top and with quite regular spacing between the islands [Fig. 1(c)]. The underlying self-organization process might become useful for the fabrication of arrays of nanometer-scale magnets. Between the 3D islands, there is still a monolayer of iron left on the W(110) substrate as is well known from complementary surface analysis studies. In Fig. 2(c) we compare the local tunneling spectra as measured above the 3D iron islands and above the iron monolayer in between. While the iron monolayer exhibits a spectrum similar to that of 2D iron islands [Fig. 2(b)] with an empty-state feature at $U = +0.2\ \text{V}$, the spectrum as measured above the 3D iron islands is markedly different: There is no empty-state feature visible at positive sample bias polarity. In contrast, a peak at $U = -0.5\ \text{V}$ shows up, corresponding to an occupied electronic state that is well known from photoelectron spectroscopy measurements. Obviously, the local electronic structure changes drastically as can be seen in the comparison of 2D and 3D iron islands.

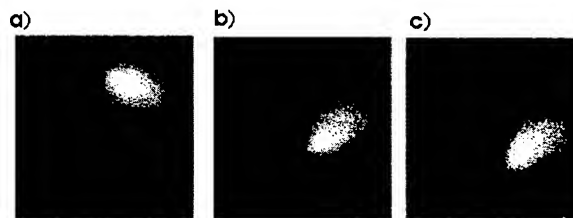


FIG. 4. Series of MFM images of a single Co dot on a GaAs substrate. Scan size: $0.5\ \mu\text{m} \times 0.5\ \mu\text{m}$. (a) Single Co dot without applied magnetic field. (b) An additional *in situ* magnetic field (100 Oe) was applied in-plane of the sample (x direction). (c) The Co dot without field (after switching off the magnetic field).

IV. MFM STUDIES OF MAGNETIC NANOSTRUCTURES

The possibility of producing large arrays of very small ferromagnetic particles by nanolithographic techniques initiated research activities that focused on their interesting magnetic properties and their possible applications as data storage media. We have fabricated Co dot arrays by electron-beam lithography and lift-off techniques on GaAs substrates with a 3 nm Cr layer to improve adhesion. In order to obtain information about the stray magnetic field of the periodically arranged Co dots, which had a thickness of 4–20 nm and a diameter of 200–250 nm, MFM studies were performed. We used a MFM instrument capable of applying a homogeneous external magnetic field during the measurements to examine the micromagnetic properties of the Co dot arrays.¹³ Magnetic force gradients were measured by detecting phase shifts in the cantilever oscillation caused by attractive or repulsive forces acting on the ferromagnetically coated tip. Figure 3 shows a MFM image of a Co dot array with a dot thickness of 17 nm and a dot diameter of 200 nm. The sample was magnetized in an external magnetic field ($H = 3000\ \text{Oe}$) applied along the x direction in-plane of the sample prior to the measurement. The MFM image shows a uniform remanent magnetization state of the Co dot array. The individual Co dots were found to be in a single domain state with north and south poles (dark and bright regions in the MFM image) at opposite edges of the particles. Figure 4 presents a series of high magnification MFM images of a single Co dot with a thickness of 7 nm. This Co dot was first scanned several times without applying any additional external magnetic field [Fig. 4(a)]. The strength of the stray field of the magnetic probe tip proved to be insufficient to affect the magnetization

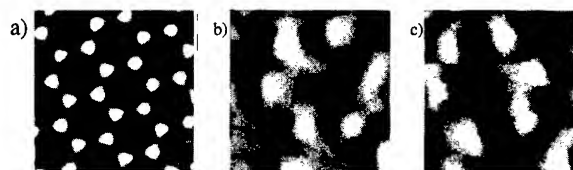


FIG. 5. (a) Topographic image ($1\ \mu\text{m} \times 1\ \mu\text{m}$) of an hexagonal array of Co dots (100 nm diameter) fabricated by nanosphere lithography. (b) MFM image ($570\ \text{nm} \times 570\ \text{nm}$) of the Co dots showing a uniform magnetized state of the single domain nanoparticles. (c) MFM image of the same Co dots as in (b) upon reversal of the magnetization direction.

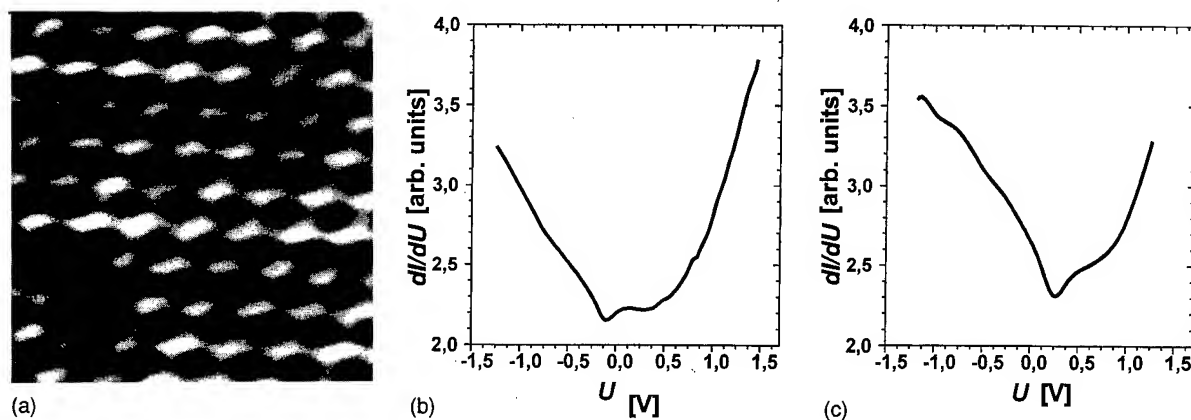


FIG. 6. (a) Atomically resolved STM image of a clean W(110) surface. Scan size: $2.1 \text{ nm} \times 2.1 \text{ nm}$. (b) Tunneling spectrum obtained with an electrochemically etched W tip on the clean W(110) surface. (c) Tunneling spectrum obtained with an *in situ* Fe-coated W tip on the clean W(110) surface.

state of the Co particle. In order to induce a switching of the magnetization of the Co dot, an external magnetic bias field was applied in-plane of the sample. The superposition of a bias field of $H = 100 \text{ Oe}$ and the stray field of the probe tip were found to lead to magnetization switching as shown in Fig. 4(b). Figure 4(c) presents the MFM image of the Co dot after switching off the external bias field. The observed magnetization direction of the Co particle is not changed with respect to that of Fig. 4(b). We found that the switching field required for changing the magnetization state of the Co dots is strongly influenced by anisotropies of the sample caused by the shape and the morphology of the micromagnetic particles. This allows optimization of the switching field that is required for local modification of the magnetic state of the individual Co dots.

Alternatively, hexagonal arrays of Co dots were fabricated by nanosphere lithography^{14,15} on mica substrates. Figure 5(a) shows the topography of an array of Co dots with a thickness of 17 nm and a diameter of about 100 nm . In Fig. 5(b) a MFM image in which the Co dots are again in a single domain state is shown. Upon magnetizing the Co dots in the opposite direction in a field of 400 Oe the MFM contrast

reverses as expected [Fig. 5(c)]. The switching field was found to be considerably lower for the smaller Co dots which makes it difficult to image their micromagnetic structure nondestructively by MFM.

V. STM AND TUNNELING SPECTROSCOPY STUDIES WITH MAGNETIC PROBE TIPS

As the size of magnetic nanostructures is reduced to 10 nm or even below, it becomes impossible to characterize their magnetic properties by MFM due to the limited spatial resolution and sensitivity of this technique. In contrast, STM/STS studies with magnetic probe tips can yield information about spin-resolved electronic and magnetic properties of surfaces down to the atomic level.^{16–19} In the present work, we have performed a comparative spectroscopic study of magnetic nanostructures by using nonmagnetic tungsten tips and ferromagnetic (Fe-coated) probe tips. To exclude spectroscopic differences of nonmagnetic origin between different tips we first performed measurements on a nonmagnetic W(110) substrate [Fig. 6(a)]. Figures 6(b) and 6(c) show two spectra obtained with an electrochemically etched W tip and

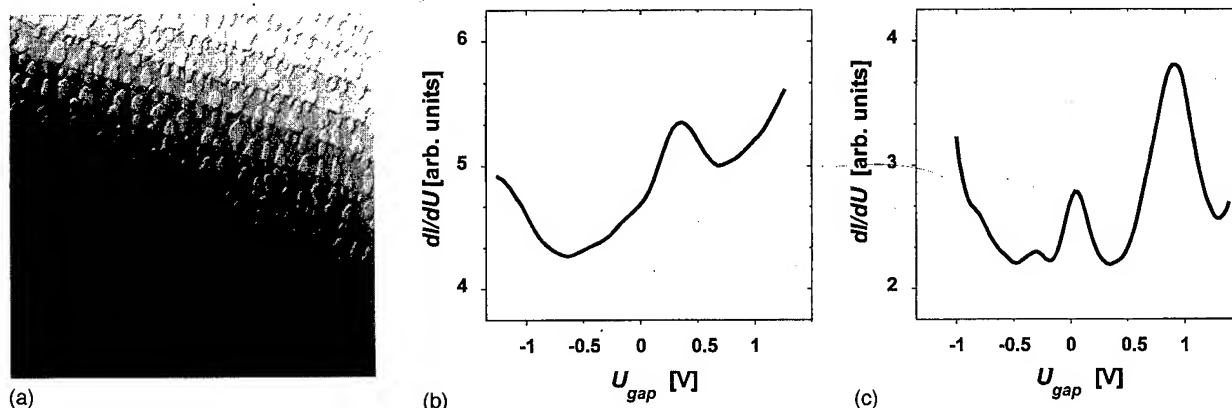


FIG. 7. (a) STM image showing 2-ML-high Fe islands on an Fe-covered W(110) surface. Scan size: $335 \text{ nm} \times 335 \text{ nm}$. (b) Tunneling spectrum obtained with an electrochemically etched W tip on a 2-ML-high Fe island. (c) Tunneling spectrum obtained with an Fe-coated W tip on a 2-ML-high Fe island.

an *in situ* Fe-coated tip, respectively. It is remarkable how similar these spectra are despite the fact that the microstructure of the tips might be completely different. Similar results were obtained with other tips as well. Subsequently, we performed a comparative spectroscopic study of 2-ML-high Fe islands [Fig. 7(a)]. Figure 7(b) shows the typical tunneling spectrum obtained by using a tungsten tip whereas Fig. 7(c) shows the corresponding spectrum obtained with Fe-coated probe tips. It is apparent, that there are dramatic differences in the two spectra. These differences might be explained by the magnetic exchange interaction across the tunneling barrier in the case of two ferromagnetic electrodes. Theoretical work to explain these experimental observations is in progress.

VI. CONCLUSIONS

Magnetic-sensitive scanning probe methods were applied to study structural, electronic, and magnetic properties of magnetic nanostructures prepared by thin film deposition under UHV conditions, by electron-beam lithography, as well as by nanosphere lithography. It was shown that drastic changes in the electronic structure occur during the transition from 0D adsorbates to 3D bulklike islands. Micromagnetic properties including domain structure and switching fields for submicron-scale Co particles were studied by MFM. The first STM spectroscopic studies of magnetic nanostructures performed with *in situ* Fe-coated tips revealed drastic differences compared to spectra obtained with nonmagnetic tungsten tips. These spectroscopic differences may prove useful for magnetic imaging of nanostructures down to the atomic scale.

ACKNOWLEDGMENTS

The authors would like to thank M. Winzer and N. Dix for preparation of the Co dot arrays by nanosphere lithogra-

phy. Financial support from the Deutsche Forschungsgemeinschaft (Grant No. Wi 1277/3-1), the Bundesministerium für Forschung und Technologie (Grant Nos. 13N6500 and 03N1023A), and the European Community (Grant No. BE 7495-93) is gratefully acknowledged.

¹P. Grünberg, R. Schreiber, Y. Pang, M. B. Brodsky, and H. Sowers, *Phys. Rev. Lett.* **57**, 2442 (1986).

²R. M. H. New, R. F. W. Pease, and R. L. White, *J. Vac. Sci. Technol. B* **13**, 1089 (1995); P. D. Ye, D. Weiss, K. von Klitzing, and K. Eberl, *Appl. Phys. Lett.* **67**, 1441 (1995); P. D. Ye, D. Weiss, R. R. Gerhardts, M. Seeger, K. von Klitzing, K. Eberl, and H. Nickel, *Phys. Rev. Lett.* **15**, 3013 (1995).

³S. Y. Chou, M. Wei, P. R. Krauss, and P. B. Fischer, *J. Vac. Sci. Technol. B* **12**, 3695 (1994); S. Y. Chou, P. R. Krauss, and P. J. Renstrom, *Appl. Phys. Lett.* **67**, 3114 (1995).

⁴Micro-STM from OMICRON, Taunusstein, Germany.

⁵Nanoscope III from Digital Instruments, CA.

⁶A. Wadas, M. Dreyer, M. Löhndorf, and R. Wiesendanger *Appl. Phys. A* **64**, 353 (1997).

⁷Noncontact UHV-AFM from OMICRON, Taunusstein, Germany.

⁸Nanosensors, Aidlingen, Germany.

⁹M. F. Crommie, C. P. Lutz, and D. M. Eigler, *Phys. Rev. B* **48**, 2851 (1993).

¹⁰M. F. Crommie, C. P. Lutz, and D. M. Eigler, *Science* **262**, 218 (1993).

¹¹M. Bode, R. Pascal, and R. Wiesendanger, *Z. Phys. B* **99**, 143 (1996).

¹²M. Bode, R. Pascal, and R. Wiesendanger, *Appl. Phys. A* **62**, 571 (1996).

¹³M. Löhndorf, A. Wadas, G. Lütjering, D. Weiss, and R. Wiesendanger, *Z. Phys. B* **101**, 1 (1996).

¹⁴J. C. Hulteen and R. P. Van Duyne, *J. Vac. Sci. Technol. A* **13**, 1553 (1995).

¹⁵M. Winzer, M. Kleiber, N. Dix, and R. Wiesendanger, *Appl. Phys. A* **63**, 617 (1996).

¹⁶R. Wiesendanger, H.-J. Güntherodt, G. Güntherodt, G. Güntherodt, R. J. Gambino, and R. Ruf, *Phys. Rev. Lett.* **65**, 247 (1990).

¹⁷R. Wiesendanger, I. V. Shvets, D. Bürgler, G. Tarrach, H.-J. Güntherodt, and J. M. D. Coey, *Europhys. Lett.* **19**, 141 (1992).

¹⁸R. Wiesendanger, *J. Magn. Soc. Jpn.* **18**, 4 (1994).

¹⁹R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy: Methods and Applications* (Cambridge University Press, Cambridge, 1994).

Magnetotransport through mesoscopic antidot arrays

Hongqi Xu^{a),b)}

Department of Solid State Physics, Lund University, Box 118, S-221 00 Lund, Sweden

(Received 8 September 1996; accepted 12 May 1997)

We present a theoretical investigation of the magnetotransport in finite antidot lattices. We have calculated the magnetoconductance, which shows slow and rapid oscillations. To interpret the rapid oscillations, we have studied the magnetoconductance correlation function and the level-spacing distribution of the systems. The results allow us to interpret the rapid oscillations as an effect of the interference of electrons traveling along chaotic trajectories in the antidot lattices. The slow oscillations can be understood in terms of pinned electron orbits around groups of antidots.
© 1997 American Vacuum Society. [S0734-211X(97)13004-9]

I. INTRODUCTION

Recently, considerable effort has been directed towards the investigation of electron transport in lateral superlattices defined by periodic modulation potentials on the plane of a two-dimensional electron gas (2DEG). Various interesting phenomena have been observed in these systems. At a weak modulation, the magnetoresistance exhibits oscillations at low magnetic-fields B .¹ These oscillations are periodic in $1/B$ and have been attributed to the oscillatory property of the Landau bandwidth as a function of B .¹ A periodically modulated 2DEG is called an antidot lattice when the modulation potential is strongly repulsive. The quenching of the Hall effect, localized magnetoresistance peaks, and regular fine oscillations around the localized peaks were observed at low magnetic fields in this system.² These phenomena have been understood in terms of classical or semiclassical theory. Irregular but reproducible magnetoresistance oscillations were also observed in a phase-coherent antidot system.³ These oscillations may have to be understood in terms of quantum-mechanical calculations.

In this work, we report on a full quantum-mechanical study of magnetotransport in phase-coherent antidot systems based on a scattering-matrix formalism. A rich structure is predicted for the mesoscopic systems. In agreement with the experimental observation of Ref. 3, we find that electron transport at low magnetic fields is characterized by slow and rapid magnetoconductance oscillations. We calculate the conductance correlation function and the correlation field B_c . We find that $B_c \cdot A \sim h/e$, where A is the total area of the antidot lattices. This result implies that the rapid oscillations result from coherent interference over entire antidot lattices. In order to gain further insight to the quantum effect, we also calculate the magnetoelectric energy-band structure for the corresponding infinite antidot lattices. We find a complex structure of the energy band in the energy range of interests to the experiments. We evaluate the nearest-neighbor band-spacing distribution function and find that the distribution function obeys the universal prediction for quantum chaos

derived from the theory of random matrices. This result is in consistence with the above interpretation for rapid magnetoconductance oscillations.

II. THEORETICAL METHOD

We consider an $N_x \times N_y$ square antidot lattice implanted with a period a in a wide ballistic channel of width w . The antidot lattice is connected to the two reservoirs modeled by two semi-infinitely long leads with the same width as the ballistic channel. We assume that the phase-breaking processes occur only inside the two reservoirs. Thus, an electron injected from, say, the left reservoir into the antidot lattice can only be elastically scattered by the antidots and the boundaries of the ballistic channel. In order to avoid artificial effects, we describe the potential of the antidot lattice by a realistic model of the form

$$V(x, y) = V_0 \{ \cos[\pi x/a] \cos[\pi(y + y_0)/a] \}^{2\beta}, \quad (1)$$

where y_0 determines the position of the antidot lattice in the channel, while V_0 and integral β control the strength and steepness of the antidot potential. With this potential, the Schrödinger equation, in general, cannot be solved analytically. In this work, we solve it numerically using the scattering-matrix technique as described in Ref. 4. This is done by dividing the antidot lattice along the transport direction (x direction) into transverse strips. We make the number of the strips, N , sufficiently large so that all the strips are thin enough to allow us to approximate the potential in each strip i by an x -independent function of the form

$$V_i(y) = V_0 \{ \cos[\pi x_i/a] \cos[\pi(y + y_0)/a] \}^{2\beta}, \quad (2)$$

where $i = 1, 2, \dots, N$ and x_i may be chosen to be the x coordinate at the center of the strip. Thus, the Schrödinger equation of motion of an electron with energy ϵ in a uniform magnetic field $\mathbf{B} = (0, 0, B)$ may be solved for each strip i by the separation of variables of the form

$$\psi^i(x, y) = e^{ik_x x} \varphi^i(y), \quad (3)$$

where $\varphi^i(y)$ satisfies the reduced one-dimensional Schrödinger equation

^{a)}Part of this work was carried out while the author was a visiting scientist at the Laboratory of Semiconductor Materials Sciences, Institute of Semiconductors, Chinese Academy of Sciences, China.

^{b)}Electronic mail: Hongqi.Xu@ftf.lth.se

$$\left\{ -\frac{\hbar^2}{2m^*} \frac{d^2}{dy^2} + \frac{m^*}{2} \omega_c^2 (y - l_B^2 k^i)^2 + V_c(y) + V_i(y) \right\} \varphi^i(y) = \epsilon \varphi^i(y). \quad (4)$$

Here, $\omega_c \equiv eB/m^*$ is the cyclotron frequency, $l_B \equiv (\hbar/eB)^{1/2}$ is the magnetic length, and $V_c(y)$ represents the confining potential of the wide quantum channel. For electron transport calculations, Eq. (4) is solved at a given electron energy ϵ for eigenwave numbers $\{k_\alpha^i\}$ and eigenwave functions $\{\varphi_\alpha^i(y)\}$. Since the corresponding operator in the eigenvalue equation for $\{k_\alpha^i\}$ is not Hermitian, k_α^i can be real, imaginary, or complex. In the calculations, we choose to normalize our eigenwave functions $\varphi_\alpha^i(y)$ so that

$$\langle \varphi_\alpha^i(y) | \varphi_\alpha^i(y) \rangle = 1, \quad (5)$$

and to define a velocity matrix $\mathbf{v}^i \equiv \{v_{\alpha\beta}^i\}$ by

$$v_{\alpha\beta}^i = \frac{\hbar}{2m^*} \int \varphi_\alpha^i(y) \left(k_\alpha^i + k_\beta^i - \frac{2y}{l_B^2} \right) \varphi_\beta^i(y) dy = v_{\alpha\beta}^i \delta_{\alpha\beta}, \quad (6)$$

where v_α^i is a short notation for $v_{\alpha\alpha}^i$ and is known as the quantum mean velocity when it is real.

The eigensolutions $\{\psi_\alpha^i(x, y)\}$ in strip region i are obtained by inserting the eigenwave numbers $\{k_\alpha^i\}$ and the eigenwave functions $\{\varphi_\alpha^i(y)\}$ into Eq. (3). The transport property of these eigensolutions is determined by their wave numbers $\{k_\alpha^i\}$. Only those eigensolutions whose wave numbers are real represent propagating modes, while the others correspond to evanescent or exploding modes. In order to overcome the difficulty in numerical calculations due to the presence of the exploding modes, we divide our eigensolutions into two groups and adopt the scattering-matrix formalism. The first group, in which the eigensolutions are denoted by $\{\psi_\alpha^{i+}(x, y)\}$, consists of those modes that are propagating forwards or evanescent, while the second group, in which the eigensolutions are denoted by $\{\psi_\alpha^{i-}(x, y)\}$, consists of all the rest, i.e., those modes that are propagating backwards or exploding.

In terms of these notations, the wave function $\Psi^i(x, y)$ of an electron with energy ϵ in strip region i can be written as

$$\Psi^i(x, y) = \sum_\alpha [a_\alpha^{i+} \psi_\alpha^{i+}(x - x_0^i, y) + a_\alpha^{i-} \psi_\alpha^{i-}(x - x_0^i, y)], \quad (7)$$

where x_0^i is the reference coordinate along the x direction for the strip region i .

The wave function in strip regions i and $i+1$ can be connected via a transfer matrix $\mathbf{M}(i, i+1)$,

$$\begin{pmatrix} \mathbf{A}_+^i \\ \mathbf{A}_-^i \end{pmatrix} = \mathbf{M}(i, i+1) \begin{pmatrix} \mathbf{A}_+^{i+1} \\ \mathbf{A}_-^{i+1} \end{pmatrix}, \quad (8)$$

where \mathbf{A}_+^i and \mathbf{A}_-^i are coefficient vectors containing $\{a_\alpha^{i+}\}$ and $\{a_\alpha^{i-}\}$, respectively. The derivation of $\mathbf{M}(i, i+1)$ is elementary and can be found in Ref. 4. The wave function in the left and right leads, $\Psi^L(x, y)$ and $\Psi^R(x, y)$, are connected via

$$\begin{pmatrix} \mathbf{A}_+^L \\ \mathbf{A}_-^L \end{pmatrix} = \mathbf{M}(L, R) \begin{pmatrix} \mathbf{A}_+^R \\ \mathbf{A}_-^R \end{pmatrix}, \quad (9)$$

where \mathbf{A}_+^L and \mathbf{A}_-^L are the coefficient vectors containing $\{a_\alpha^{L+}\}$ and $\{a_\alpha^{L-}\}$, \mathbf{A}_+^R and \mathbf{A}_-^R are the coefficient vectors containing $\{a_\alpha^{R+}\}$ and $\{a_\alpha^{R-}\}$, and $\mathbf{M}(L, R)$ is the total transfer matrix of the system given by $\mathbf{M}(L, R) = \mathbf{M}(L, 1) \cdot \mathbf{M}(1, 2) \cdots \mathbf{M}(N-1, N) \cdot \mathbf{M}(N, R)$.

It is well known that in the calculations for the total transfer matrix $\mathbf{M}(L, R)$, one often suffers a numerical instability due to the presence of both exponentially growing and exponentially decaying terms in the formulation. This problem may be removed by rewriting Eq. (9) as

$$\begin{pmatrix} \mathbf{A}_+^R \\ \mathbf{A}_-^R \end{pmatrix} = \mathbf{S}(L, R) \begin{pmatrix} \mathbf{A}_+^L \\ \mathbf{A}_-^L \end{pmatrix}, \quad (10)$$

where $\mathbf{S}(L, R)$ is known as the scattering matrix of the system and can be obtained iteratively with the help of the transfer matrix $\mathbf{M}(i, i+1)$. For details of this procedure, we refer to our earlier works in Ref. 4.

In the calculations for electron transport, the boundary condition imposed on the wave function of the electron is $\mathbf{A}_+^L = \mathbf{I}_\gamma$ and $\mathbf{A}_-^R = 0$, where \mathbf{I}_γ is a unit vector with elements given by $(\mathbf{I}_\gamma)_\alpha = \delta_{\alpha\gamma}$. The wave-function coefficients \mathbf{A}_+^R and \mathbf{A}_-^L are simply obtained by inserting this boundary condition into Eq. (10). The magnetoconductance of the system is then calculated in the linear-response regime using

$$\begin{aligned} G &= \frac{2e^2}{h} \sum_\gamma^{(R)} \left(1 + \sum_\alpha^{(R)} \frac{v_\alpha^{L-}}{v_\gamma^{L+}} |a_\alpha^{L-}|^2 \right) \\ &= \frac{2e^2}{h} \sum_\gamma^{(R)} \sum_\alpha^{(R)} \frac{v_\alpha^{R+}}{v_\gamma^{L+}} |a_\alpha^{R+}|^2, \end{aligned} \quad (11)$$

where (R) indicates that the summations are taken over those values of γ and α for which v_γ^{L+} , v_α^{L-} , and v_α^{R+} are real. For further details of the calculations, we refer to Ref. 4.

III. RESULTS AND DISCUSSION

Figure 1 shows the magnetoconductance calculated for two antidot lattices with period $a = 100$ nm, $V_0 = 1$ eV, and $\beta = 10$. The two antidot lattices have 20 periods along the transport x direction. Along the transverse y direction, they have three and four periods, respectively. In the calculations, we have assumed an electron effective mass $m^* = 0.067m_e$ and the Fermi energy $E_F = 14$ meV. The calculated magnetoconductance shows a generic structure similar to the experimental observation of Ref. 3: $G(B)$ exhibits slow and rapid oscillations. The slow oscillations were interpreted as an effect originating from pinned classical electron orbits around groups of antidots.³ Slow and rapid oscillations have also been obtained in the conductance calculated as a function of the Fermi energy at zero magnetic field.⁵ The slow oscillations in the zero- B case were interpreted in terms of the Bragg reflection of the electron waves by the periodic potential.⁵

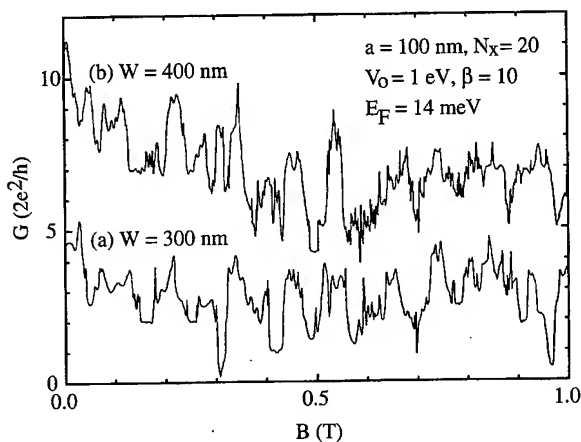


FIG. 1. Calculated magnetoconductance at $E_F = 14$ meV for antidot lattices with $a = 100$ nm, $N_x = 20$, $V_0 = 1$ eV, and $\beta = 10$. Curve (a) is for $w = 300$ nm ($N_y = 3$), while curve (b) is for $w = 400$ nm ($N_y = 4$).

The rapid magnetoconductance oscillations cannot be simply interpreted in terms of classical mechanics. Their origin is purely quantum mechanical. To understand these oscillations, we calculate the magnetoconductance correlation function defined by

$$F(\Delta B) = \langle [G(B) - \langle G(B) \rangle] \times [G(B + \Delta B) - \langle G(B + \Delta B) \rangle] \rangle. \quad (12)$$

The correlation field ΔB_c is defined as $F(\Delta B_c) = F(0)/2$. Figure 2 shows the calculated magnetoconductance correlation function for the 20×4 antidot lattice from the magnetoconductance curve as shown in Fig. 1(b). The determined correlation field ΔB_c is about 0.009 T. If the total area of the antidot lattice $A = 8 \times 10^5$ nm² is used, we obtained $\Delta B_c \cdot A \sim h/e$. This result implies that the rapid oscillations have originated from the coherent interference of electron waves traveling over the entire antidot lattice.

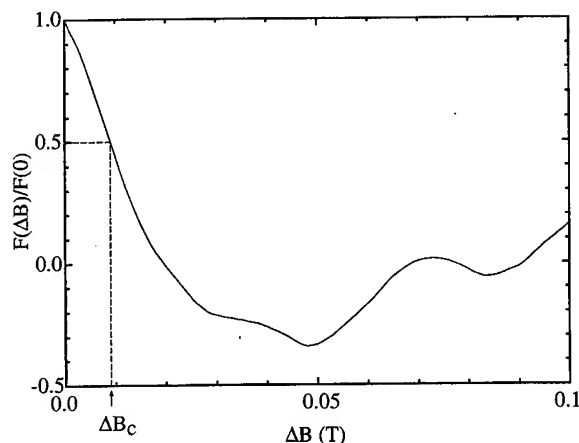


FIG. 2. Normalized correlation function of the magnetoconductance oscillations of the 20×4 antidot lattice, as shown in curve (b) of Fig. 1. The determined corresponding correlation field ΔB_c is indicated by a vertical arrow.

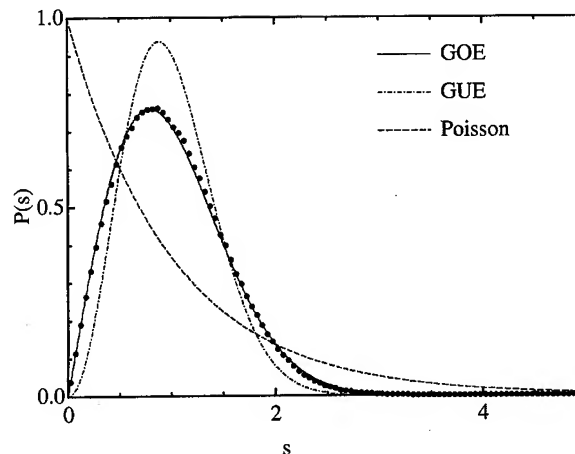


FIG. 3. Energy-band-spacing distribution function of the quasi-two-dimensional antidot crystal with $N_x = \infty$ and $N_y = 5$ ($w = 500$ nm). Other parameters were taken to be the same as in Fig. 1. The statistical average was taken over 300 magnetolectric energy bands and 40 magnetic-field values in the range of 0.1–0.3 T.

The understanding of the magnetotransport property in finite antidot lattices can also be obtained from the study of the energy spectra of the systems. Since the quantum systems are difficult to solve, we instead study the magnetolectric energy-band structure for the corresponding antidot lattices that are infinite along the x direction. We find an extremely complex band structure that reminds us about quantum chaos. We calculate the band-spacing distribution function $P(s)$ and find that $P(s)$ is close to the Gaussian orthogonal ensemble (GOE) distribution derived from the theory of random matrices. This result, as shown in Fig. 3, is a clear evidence for the presence of chaos in the antidot systems. The rapid oscillations can, therefore, be interpreted as due to the interference of electrons traveling along chaotic trajectories in the antidot lattices.

ACKNOWLEDGMENTS

The author acknowledges the hospitality of the Laboratory of Semiconductor Materials Sciences, Institute of Semiconductors, Chinese Academy of Sciences. This work was supported by the Swedish Natural Science Research Council (NFR), the Swedish National Board for Industrial and Technical Development (NUTEK), and the Swedish Research Council for Engineering Sciences (TFR).

¹D. Weiss, K. V. Klitzing, K. Ploog, and G. Weimann, *Europhys. Lett.* **8**, 179 (1989); R. R. Gerhardt, D. Weiss, and K. V. Klitzing, *Phys. Rev. Lett.* **62**, 1173 (1989); R. W. Winkler and J. P. Kotthaus, *ibid.* **62**, 1177 (1989).

²D. Weiss, M. L. Roukes, A. Mensching, P. Grambow, K. V. Klitzing, and G. Weimann, *Phys. Rev. Lett.* **66**, 2790 (1991); D. Weiss, K. Richter, A. Mensching, R. Bergman, H. Schweizer, K. V. Klitzing, and G. Weimann, *ibid.* **70**, 4118 (1993).

³R. Schuster, K. Ensslin, D. Wharam, S. Kühn, J. P. Kotthaus, G. Böhm, W. Klein, G. Tränkle, and G. Weimann, *Phys. Rev. B* **49**, 8510 (1994).

⁴Hongqi Xu, *Phys. Rev. B* **50**, 8469 (1994); **52**, 5803 (1995).

⁵Hongqi Xu, *Phys. Rev. B* **50**, 12 254 (1994).

Magnetic property of NiO ultrafine particles with a small Ni core

Sanshiro Sako^{a)}

Department of Physics, Faculty of Education, Mie University, Tsu 514, Japan

Kazunari Ohshima

Laboratory of Physics, Matsumoto Dental College, Shiojiri 399-07, Japan

Masahiro Sakai and Shunji Bando

Instrument Center, Institute for Molecular Science, Myodaiji, Okazaki 444, Japan

(Received 8 September 1996; accepted 8 March 1997)

The magnetic nature of a NiO shell and a Ni core were separately investigated in the temperature dependence of the magnetization curve of the samples with the average sizes of 1.6, 2.0, 2.2, 2.5, 2.8, 3.0 and 3.2 nm. The superparamagnetism of the NiO shell and Ni core were observed in specific temperature ranges. The increase of the magnetic moment per atom in the metallic free cluster with decreasing size were not observed in the small Ni core with the NiO shell though the proper saturation magnetization of Ni core was estimated. © 1997 American Vacuum Society. [S0734-211X(97)05104-4]

I. INTRODUCTION

In systems with interactions among atoms, there exist differences in the physical properties of atoms on the surface and atoms in a particle because of the difference of coordination number. In the magnetic properties of a transition metal cluster or ultrafine particle, the characteristic nature of the surface atoms is expected to dominate¹⁻⁴ because the ratio of the surface to volume is large. In cluster beam experiments,^{5,6} the important fact that the magnetic moment per atom increases as the size of a particle decreases was confirmed. Also in magnetic compounds, the magnetic moment per atom on the surface should be different from that of inner atoms at temperatures lower than the Néel temperature or Curie temperature because the spin ordering is due to the interatomic exchange interaction. Therefore, a magnetic nature^{7,8} different from that of the bulk naturally should appear in the cluster or ultrafine particle of a magnetic compound.

We previously reported the particle size dependence of the Néel temperature and other magnetic properties of ultrafine particles of the magnetic oxides MnO (Ref. 9) and CoO (Ref. 10). In this article, as a part of a study of transition metal oxides, we report experimental results on the NiO ultrafine particle with crystal structure and magnetic nature analogous to the previous oxides. A stable, small metallic core in a NiO particle remained in air, as was true in the case of the CoO particle. The size of the core was estimated from x-ray photoelectron spectroscopy (ESCA) measurement. As a result, we could estimate the saturation magnetization of the core Ni, but we could not obtain the increase of magnetic moment per atom with decreasing size. Furthermore, we found that the exchange anisotropy through the interface in the case of Ni particle is small in comparison with the cases of Co and Fe particles.

II. EXPERIMENT AND RESULTS

Ultrafine particles were produced using the initial island structure resulting from vacuum evaporation.¹¹ Five layers of Ni particles were evaporated on a 25- μ m-thick Mylar film by alternative evaporation with LiF (20 nm) which was used as the matrix substance. Those sandwiched Ni particles oxidized in air. After oxidation for several hours a stable small Ni core remained in a NiO particle. That was confirmed by the observation of a change of the electron diffraction pattern with time. The average size (diameter) of particles was controlled by the deposition controller (Inficon, IC6000). We produced seven samples with the average sizes of 1.6, 2.0, 2.2, 2.5, 2.8, 3.0, and 3.2 nm. Examples of a transmission electron microscopy (TEM) micrograph and an electron dif-

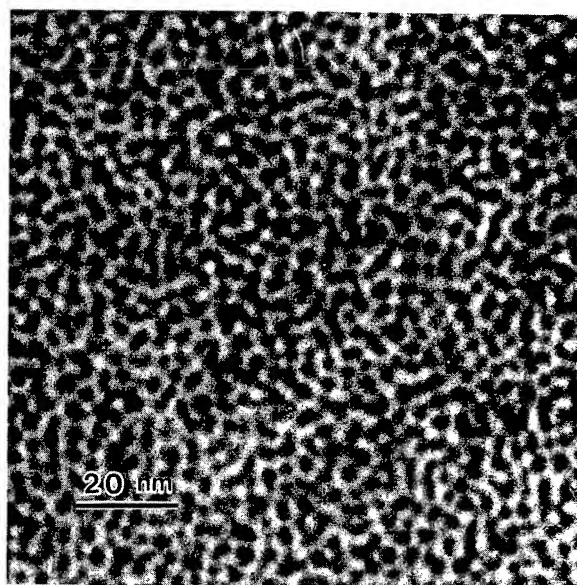


FIG. 1. The micrograph of NiO ultrafine particles with the average size of 2.8 nm.

^{a)}Electronic mail: sako@edu.mie-u.ac.jp

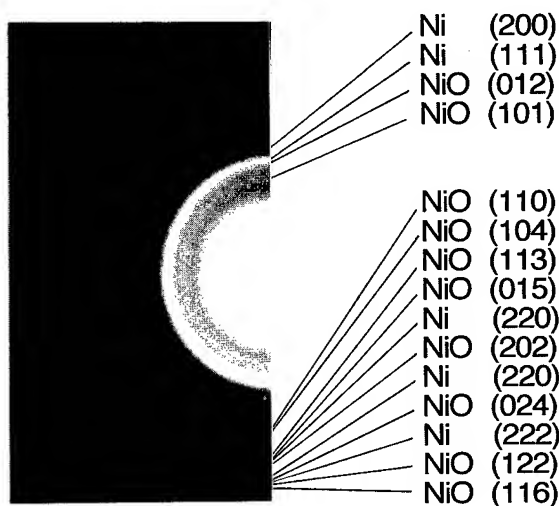


FIG. 2. The electron diffraction pattern of NiO ultrafine particles with the Ni core. The average size is 2.8 nm. Weak lines of Ni(111) and Ni(200) of fcc structure are recognized in the strong NiO lines.

fraction pattern of NiO particles are shown in Figs. 1 and 2, respectively, in the sample with the average size of 2.8 nm. The crystal structures of Ni and NiO in a particle were the same face centered cubic (fcc) and rhombohedral structures as those in the bulk state. We utilized x-ray photoelectron spectroscopy in order to estimate the quantity of the stable Ni core. An example of the ESCA spectra is shown in Fig. 3 for the sample with the average size of 2.8 nm. The spectral intensities of the Ni $2p_{3/2}$ level in the NiO shell and Ni core were compared. The ratio, α , of the number of Ni atoms in Ni core to those in NiO shell is shown in Table I. These values nearly agreed with the estimate from the line intensities of Ni and NiO in the electron diffraction patterns.

The magnetization curves were measured to investigate the magnetic nature of the NiO ultrafine particles using a high sensitivity Faraday-type magnetic balance. Measurements were made over the magnetic field range of -50 to 50 kOe in the temperature range from 2 to 250 K. Examples of the magnetization curves are shown in Figs. 4(a) and 4(b) for the sample with the size of 2.0 and 2.8 nm, respectively.

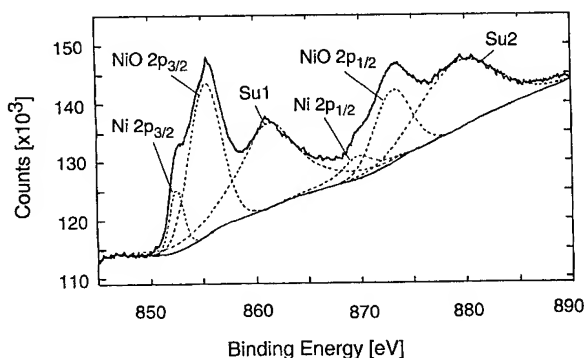
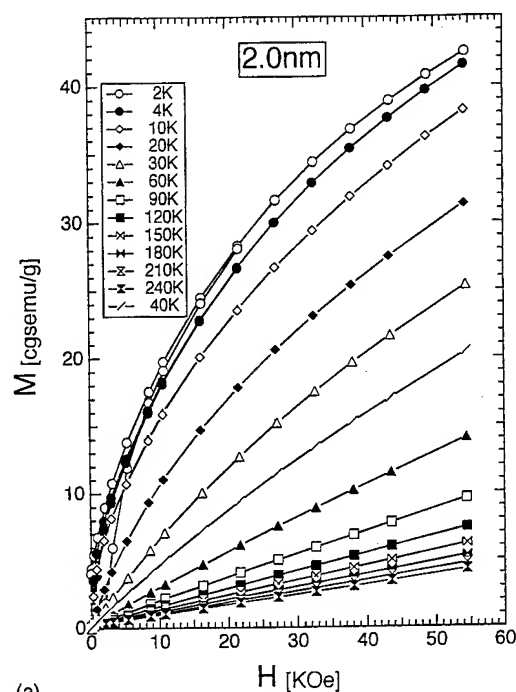


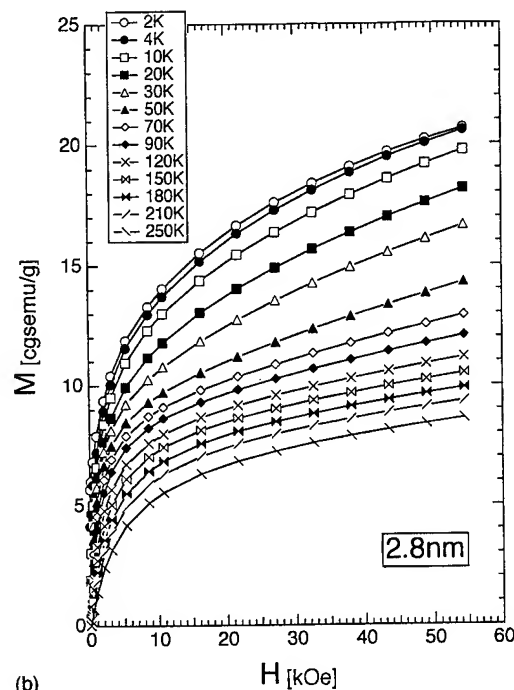
FIG. 3. The ESCA spectrum of a NiO ultrafine particle with a Ni core. The average size is 2.8 nm.

TABLE I. α is the ratio of the number of Ni atoms in Ni core to those in NiO shell for samples with various average size.

Size of sample (nm)	α	Size of core (nm)
3.2	0.25 ± 0.07	1.73 ± 0.20
3.0	0.20 ± 0.06	1.61 ± 0.08
2.8	0.17 ± 0.05	1.34 ± 0.10
2.5	0.12 ± 0.03	1.06 ± 0.08
2.2	0.07 ± 0.01	0.81 ± 0.05



(a)



(b)

FIG. 4. The magnetization curve at various temperatures for samples with average size of (a) 2.0 nm and (b) 2.8 nm.

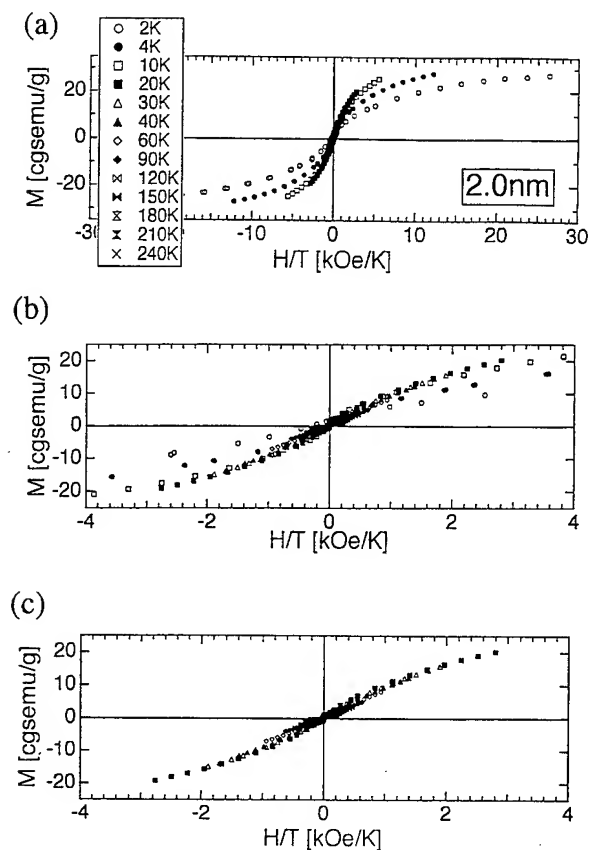


FIG. 5. The rearrangement of Fig. 4(a). (a) The magnetization curves plotted against H/T . (b) The magnetization curves near zero field in (a) are enlarged. (c) The magnetization curves at 2, 4, and 10 K in (b) are deleted.

When the unit of the horizontal axis is converted to H/T , Fig. 4(a) changes into Fig. 5(a). Figure 5(b) is the extension of the neighborhood of zero field in Fig. 5(a). When the curves at 2, 4, and 10 K where the large residual magnetization appears were excluded from Fig. 5(b), Fig. 5(c) is obtained. Then, remaining curves at any temperature overlap well. The relation between Fig. 4(b) and Fig. 6 is also the same. This means that those curves obey the Langevin function. In Fig. 6, we can distinguish two components; one is the part with a large change near zero field and the other is the part with a constant gradient. The Curie temperature of the ferromagnetic bulk Ni is about 630 K and the Néel temperature of the antiferromagnetic bulk NiO is about 520 K. Such magnetic transition temperatures are expected to decrease in a small particle, but in the lower temperature range of our measurement, magnetic ordering will take place. Accordingly, we considered that above two paramagnetic components are attributed to the superparamagnetism of the Ni core and that of the NiO shell.^{12,13} The former corresponds to the part which showed the large change near zero field and the latter corresponds to the part which showed the constant gradient on the higher field because the former became stronger as the average particle size increased.

The magnetic susceptibility of the superparamagnetism of the NiO shell, which was obtained from the gradient of the

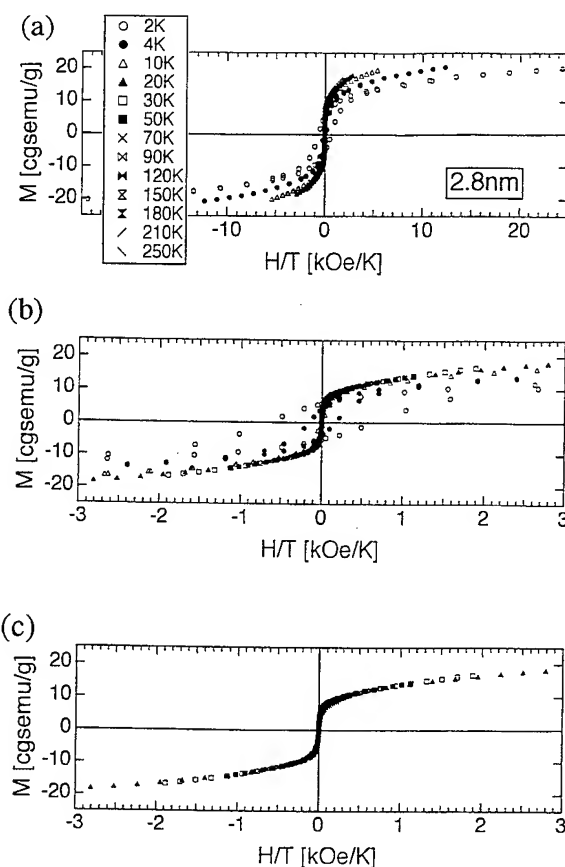


FIG. 6. The rearrangement of Fig. 4(b). (a) The magnetization curves plotted against H/T . (b) The magnetization curves near zero field in (a) are enlarged. (c) The magnetization curves at 2, 4, and 10 K in (b) are deleted.

higher magnetic field, is shown in Fig. 7 for samples with various average sizes. It increases as the average size decreases similar to other NiO particle experiments.^{13,14} According to Néel,¹² the antiferromagnetic small particle has the magnetic moment $N^{1/2}\mu$ per particle due to the statistical fluctuation, where N and μ are the number of Ni atoms in the sublattice and the magnetic moment per Ni atom, respectively. Therefore, as a group, the antiferromagnetic particles show the superparamagnetism described as $\chi = nN\mu^2/3kT$, where n is the number of particles per unit mass. Then, χ should be size independent and differ from the experimental results shown in Fig. 7 because $1/n$ and N are both proportional to the volume of a particle. The disorder of spins near surface (canting effect) had been observed¹⁵⁻¹⁷ in experiments of the Mössbauer effect. The canting effect increases as the size decreases. This may be due to the size dependence in Fig. 7.

From the extrapolation of the high field gradient of the magnetization in Fig. 4 to zero field, we can obtain the saturation magnetization M_s of the Ni core. The M_s , calibrated by the mass of Ni core estimated from Table I, is shown in Fig. 8 for samples with various sizes. In the low temperature range, M_s increased as the size decreased and the relation was reversed in higher temperature range. That is to say, the magnetic moment per atom increases as the size decreases, in

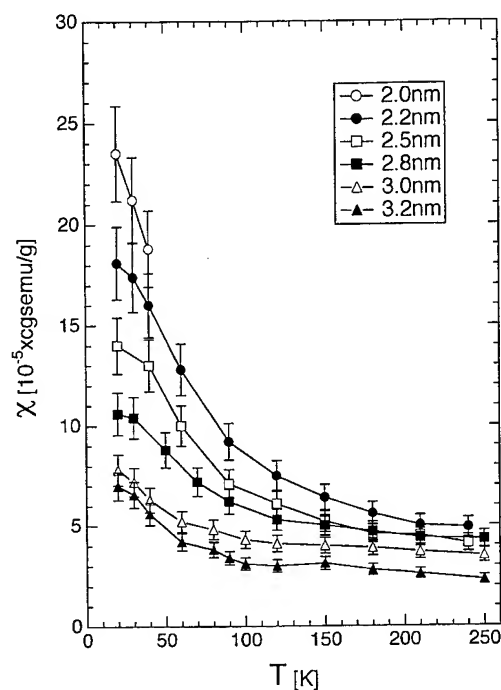


FIG. 7. The temperature dependence of the magnetic susceptibility of the NiO shell in samples with various sizes.

the low temperature range. The change of M_s with decreasing temperature is larger as the size is smaller. We think that such a large change of M_s in small particles is due to the ferrimagnetism in a nickel oxide other than NiO which is not

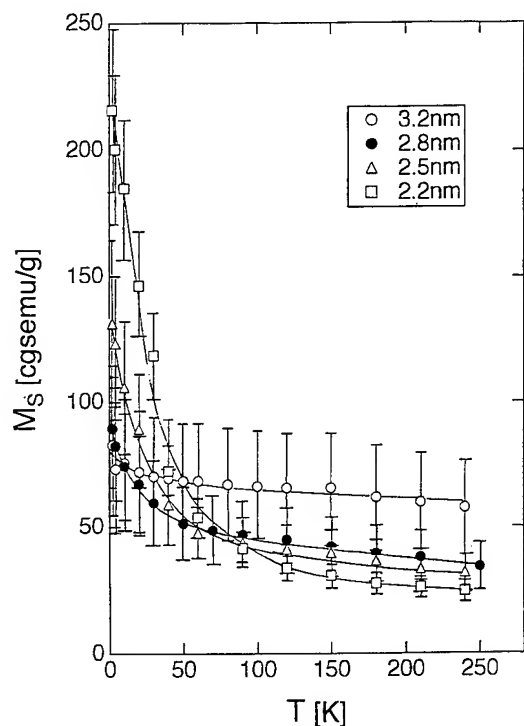


FIG. 8. The temperature dependences of the saturation magnetization M_s of the Ni core in samples with various sizes.

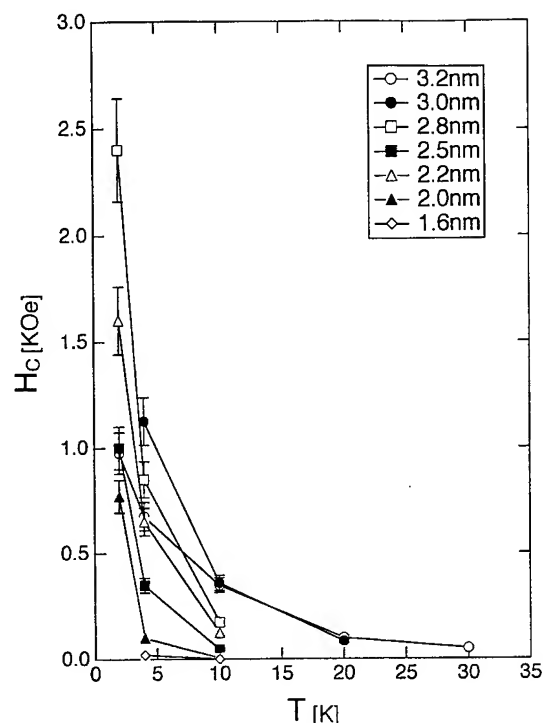


FIG. 9. The temperature dependence of coercive force in samples with various sizes.

observed by the electron diffraction, in the same manner as the case of Co particles.¹⁸ M_s at the highest temperature in Fig. 8 where NiO shows the paramagnetism is close to the intrinsic M_s of Ni core. M_s , that is, the magnetic moment per atom at this temperature decreases with decreasing size in contrast to the case of a free cluster. However, the size difference does not seem to be meaningful because it is comparable with error bar.

The temperature dependence of the coercive force obtained from the hysteresis loop is shown for samples with various average sizes in Fig. 9. The coercive force has the tendency to increase with size reflecting the difference of bound energy KV , where K and V are the anisotropy energy and the particle volume, respectively. The size dependence of the blocking temperature where H_c becomes zero can be explained by the same reasoning. In the case of Co (Ref. 18) and Fe (Ref. 19) particles, the size dependence of the coercive force was reversed in the low temperature range for the relatively stronger exchange anisotropy due to the larger ratio of oxide to core metal in smaller particles. In the case of Ni, such a reverse did not take place. It may suggest that the exchange anisotropy in the case of Ni is weaker than those in the case Co and Fe.

III. CONCLUSIONS

In conclusion, we investigated separately the magnetic nature of the NiO shell and the Ni core in a particle based on the large difference of magnetic moment. We found that NiO shells and Ni cores both showed superparamagnetism, and that the superparamagnetism was affected by the slight fer-

rimagnetic oxide or canting effect. A weak exchange anisotropy in comparison with the case of Co and Fe was expected from coercive data. We tried to estimate the intrinsic magnetic moment per atom in Ni core using the mass of only Ni core, but in our case of a metal surrounded by oxide we could not reproduce the tendency found by the free cluster beam experiment.

ACKNOWLEDGMENTS

The authors thank the Instrument Center, Institute for Molecular Science, for the use of a magnetic balance, and the Electron Microscopy Center, Mie University for the use of a transmission electron microscopy.

¹K. Lee and J. Callaway, Phys. Rev. B **48**, 15 358 (1993).

²S. N. Khanna and S. Linderth, Phys. Rev. Lett. **67**, 742 (1991).

³F. Liu, M. R. Press, S. N. Khanna, and P. Jena, Phys. Rev. B **39**, 6914 (1989).

⁴F. Liu, S. N. Khanna, and P. Jena, Phys. Rev. B **43**, 8179 (1991).

⁵J. P. Bucher, D. C. Douglass, and L. A. Bloomfield, Phys. Rev. Lett. **66**, 3052 (1991).

⁶I. M. L. Billas, A. Châtelain, and W. A. de Heer, Science **265**, 1682 (1994).

⁷F. Aguilera-Granja and J. L. Morán-López, Phys. Rev. B **31**, 7146 (1985).

⁸J. L. Morán-López and K. H. Bennemann, Solid State Commun. **88**, 101 (1993).

⁹S. Sako and K. Ohshima, J. Phys. Soc. Jpn. **64**, 944 (1995).

¹⁰S. Sako, K. Ohshima, M. Sakai, and S. Bandow, Surf. Rev. Lett. **3**, 109 (1996).

¹¹S. Sako, Jpn. J. Appl. Phys. **1** **28**, 1915 (1989).

¹²L. Néel, J. Phys. Soc. Jpn. **17**, 676 (1962).

¹³J. Cohen, K. M. Creer, R. Pauthhnet, and K. Srivastava, J. Phys. Soc. Jpn. **17**, 685 (1962).

¹⁴J. T. Richardson and W. O. Milligan, Phys. Rev. **102**, 1289 (1956).

¹⁵Q. A. Pankhurst and R. J. Pollard, Phys. Rev. Lett. **67**, 248 (1991).

¹⁶J. M. D. Coey, Phys. Rev. Lett. **27**, 1140 (1971).

¹⁷S. Gangopadhyay, G. C. Hadjipanayis, C. M. Sorensen, and K. J. Klambunde, J. Appl. Phys. **73**, 6964 (1993).

¹⁸S. Gangopadhyay, G. C. Hadjipanayis, C. M. Sorensen, and K. J. Klambunde, IEEE Trans. Magn. **28**, 3174 (1992).

¹⁹L. Yiping, G. C. Hadjipanayis, and C. M. Sorensen, J. Appl. Phys. **67**, 4502 (1990).

Magnetic force microscope images of magnetic domains in magnetic garnet

F. Tian, C. Wang, G. Y. Shang, N. X. Wang, and C. L. Bai^{a)}
STM Laboratory, Institute of Chemistry, Academia Sinica, Beijing 100080, China

(Received 12 September 1996; accepted 31 March 1997)

Magnetic domains in $(\text{YGdBi})_3(\text{GaFe})_5\text{O}_{12}$ magnetic garnet have been studied using magnetic force microscopy (MFM). MFM results have demonstrated that the undisturbed magnetic domain structure in the garnet could be obtained with a soft magnetic tip. It is further observed that the stray field emanating from a hard tip could magnetize the low-coercivity garnet and alter the widths of the domain region. In addition, when an external magnetic field is applied to the garnet, a sequence of variations of the magnetic domain configuration in the garnet is clearly identified. The discussions on the effect of magnetic tip of MFM on magnetic domains and the external field on magnetic tips will also be presented. © 1997 American Vacuum Society. [S0734-211X(97)04104-8]

I. INTRODUCTION

The domain structure in garnets were usually studied using ferrofluid decoration methods and the magneto-optic Faraday effect, which have led to great advances in this field. As research activities proceed, the demands for higher resolution is becoming increasingly strong. Soon after magnetic force microscopy (MFM) was available, it was quickly proven to be a powerful technique in imaging a wide range of magnetic materials¹⁻⁸ and applied to investigate garnet with much success.^{9,10} Due to the low coercivity in garnet and the inevitable magnetostatic coupling between tip and sample, research efforts have seen difficulties in studying the fringing field in magnetic garnets which is essential to understand the magnetic structure of the material. Therefore, the importances of the interaction between tip and sample is increasing. So far, using MFM in the presence of an external magnetic field, the approach for the studies of the switching of the magnetization of individual isolated particles,^{11,12} the measuring of the coercivity of recording media,¹³ and the imaging of the magnetic structures in recording media undergoing dc erasure,¹⁴ are also reported, which is deemed more practical for the application of MFM.

In this work, we have demonstrated that the undisturbed magnetic domain structure could be obtained using MFM with a soft magnetic tip. It is further observed that the stray field emanating from a hard tip could magnetize the garnet and alter the widths of the domain region. By choosing appropriate tip-sample separation, we could minimize the tip effect. We also report the observation of the static changes of the domain configuration in $(\text{YGdBi})_3(\text{GaFe})_5\text{O}_{12}$ garnet within an external magnetic field. Sequential variations of the magnetic domain configuration in the garnet are observed and could be associated with the strength of the applied field. In this article, the effects of the magnetic tip are presented first for the purpose of assisting to bring out the effects attributed to the external field. We have also demonstrated that the details of the magnetic configuration could be analyzed semiquantitatively.

II. EXPERIMENT

The samples under study were epitaxially grown $(\text{YGdBi})_3(\text{GaFe})_5\text{O}_{12}$ garnet films on $\text{Gd}_3\text{Ga}_5\text{O}_{12}$ substrates.¹⁵ The magnetic garnet film has a measured saturation magnetization $4M_s = 185$ G, film thickness $h = 5.0$ μm , stripe width of 5.55 μm , and a collapse field of 87 Oe. The magnetic domains of garnet were imaged by Nanoscope III DimensionTM 3000 scanning probe microscope, manufactured by Digital Instruments, Inc. The cantilever was a Nano-Probe magnetically coated tip with a 225 μm length and a resonant frequency of approximately 70 kHz. The tip was magnetized along the tip direction, which was perpendicular to the imaging surface of the sample. The topographical and magnetic images were obtained simultaneously using tapping and lift modes, respectively.

III. RESULTS AND DISCUSSION

During MFM experiments, the magnetic tip is scanned over the surface of a sample and detects the stray field from the sample, $H(r)$, which can be calculated, according to

$$H(r) = \int_{\text{sample}} \frac{3\hat{n}[\hat{n} \times m_2(r_2)] - m_2(r_2)}{r^3} dV',$$

where r is the observer point outside of the sample. $m_2(r_2)$ is the magnetization of the volume element dV' in the sample, and $\hat{n} = r/|r|$. The force acting on the tip in the direction of the z axis can be given as

$$F_z = \int_{\text{tip}} dV'' \int_{\text{sample}} dV' f_z(r_1 - r_2),$$

where $f_z(r)$ is the interaction force between two magnetic dipoles, $m_1(r_1)$ and $m_2(r_2)$, at a distance $r = r_1 - r_2$:

$$f_z(r) = \frac{\partial}{\partial z} \left(\frac{m_1(r_1) \{3\hat{n}[\hat{n} \times m_2(r_2)] - m_2(r_2)\}}{r^3} \right),$$

where $m_1(r_1)$ is the magnetization of the volume element dV'' in the tip. One simple but particularly instructive approximation is to assume that the tip consists of a point dipole, $m_1 = M$.² Thus, $F_z = M(\partial/\partial z)H(r)$.

^{a)} Author to whom all correspondence should be addressed.

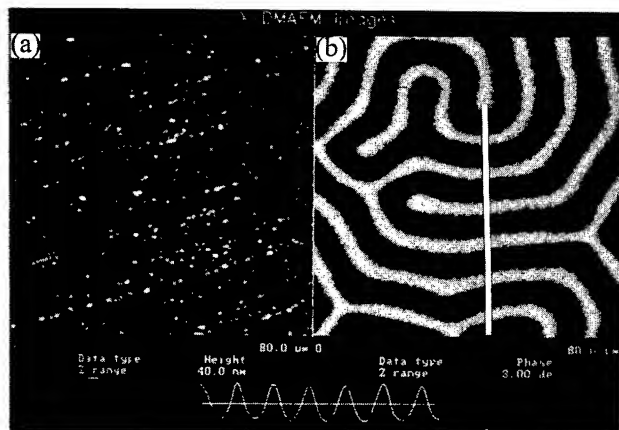


FIG. 1. (a) AFM topography of epitaxially grown $(\text{YGdBi})_3(\text{GaFe})_5\text{O}_{12}$ magnetic garnet. (b) Corresponding MFM image of the garnet taken simultaneously with a soft magnetic tip. The tip-sample distance was 200 nm. Plot of the cross-sectional profile was taken as marked in the Fig. 1(b).

Figure 1(a) shows an atomic force microscope (AFM) topography of $(\text{YGdBi})_3(\text{GaFe})_5\text{O}_{12}$ garnet. And Fig. 1(b) is the corresponding MFM image of the garnet taken simultaneously using a soft magnetic tip magnetized normal to the sample surface. In this figure, the bright and dark stripes represent magnetic domains with opposite magnetization as explained in the following paragraph. The cross-section profile, which was taken across the stripes as marked in Fig. 1(b), agreed with the numerical simulations using the above formalism (calculated results are not shown). It is apparent that the measured widths of the stripe domains are approximately equal and the contrast of the domain regions is nearly symmetric.

If the magnetization directions of magnetic tip and domains in the sample are antiparallel, the interaction between them is repulsive, giving a bright contrast in the MFM image. While in the case of parallel magnetization, the interaction is attractive, appearing as a dark contrast in image. The periodicity of the MFM image contrast in Fig. 1(b) was measured to be approximately $10.8 \mu\text{m}$, which is twice the stripe width. This is consistent with the domain widths measured employing the magneto-optic Faraday effect. This suggests that the magnetic domains observed with the soft tip is indeed an undisturbed one.

MFM images with a hard magnetic tip are also obtained. In Fig. 2, the cross-section profile changes appreciably compared with Fig. 1(b). The widths of the bright and dark stripes are unequal. The widths of the dark domains are wider than that of the bright domains. Moreover, the interaction of the bright domains with the hard tip is far stronger than that of the dark domains as evidenced by the apparent asymmetry of the line shape. This suggests that the stray field emanating from a hard tip could magnetize the garnet locally (enlarging the parallel region and reducing the antiparallel region) and should not be negligible. Figure 3 shows the magnified image of Fig. 2. In Fig. 3(b), the little domains along the bright stripe suggest that we have observed the domain wall.

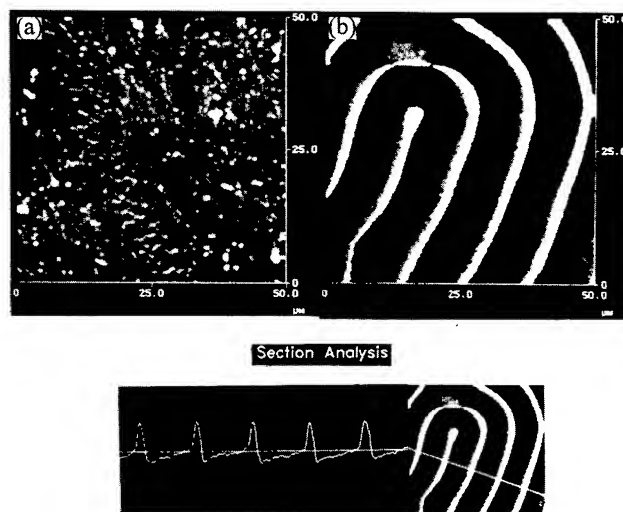


FIG. 2. The AFM (a) and MFM (b) images of $(\text{YGdBi})_3(\text{GaFe})_5\text{O}_{12}$ magnetic garnet recorded with a hard magnetic tip. The tip-sample distance was 150 nm. The cross-sectional profile was taken as marked in the figure.

The measured widths of stripe domains versus tip-sample distance demonstrates that, with the tip-sample separation increasing, the widths of bright domains gradually increase and that of the dark decrease. However, the image contrast period is nearly unchanged. The reason could be that, with the tip-sample separation increasing, the stray field from magnetic tip gradually decreases. So the effect of a hard tip on the widths of domains is expected to be gradually reduced. In addition, the observed domain configuration of the garnet remains mostly unchanged. When the tip was positioned higher than 1500 nm from the sample surface, the widths of the bright stripes returned to $5.34 \mu\text{m}$, which is close to the original width of the undisturbed domains in the garnet. The fact that the periodicity does not alter substantially suggests that the stray field from the hard tip is very small.¹⁶

When an external magnetic field (H_{ext}) was applied to the garnet, we observed sequential changes in the magnetic domains in the garnet using a hard tip. We prefer this combination since hard tips have the advantage of most likely preserving their magnetization throughout experiments.¹⁷ As

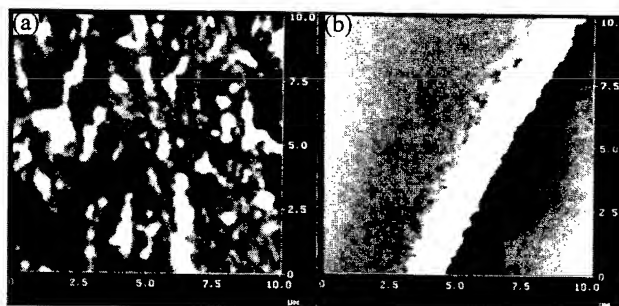


FIG. 3. The magnified AFM(a) and MFM(b) images of $(\text{YGdBi})_3(\text{GaFe})_5\text{O}_{12}$ magnetic garnet shown in Fig. 2. The tip-sample distance was 80 nm.

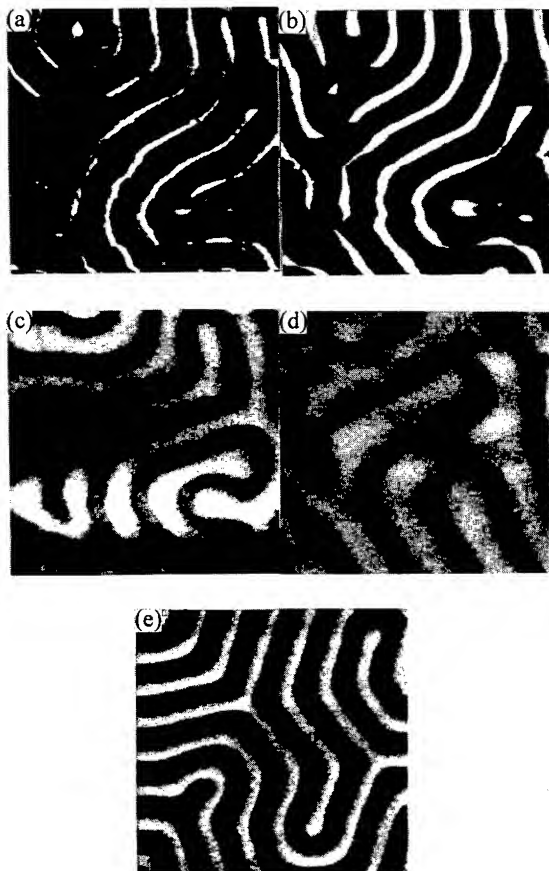


FIG. 4. The MFM image of the garnet using a hard tip with an external magnetic field. The tip-sample distance was 200 nm. The imaged region is $80\text{ }\mu\text{m} \times 80\text{ }\mu\text{m}$. The external field is antiparallel to the magnetization of magnetic tip. (a) $H_{\text{ext}}=0\text{ G}$, (b) $H_{\text{ext}}=21\text{ G}$, (c) $H_{\text{ext}}=43\text{ G}$, (d) $H_{\text{ext}}=61\text{ G}$, and (e) the MFM image of the garnet after H_{ext} is removed.

shown in the series of images in Fig. 4, it is evident that the observed domain configuration has been strongly affected by applied field. Figure 5 shows the changes of the measured widths of the stripe domains versus the external magnetic field. With the increasing H_{ext} , the widths of the bright domain increase gradually and the widths of the dark domains decrease slightly, while the domain configuration also changes appreciably [Figs. 4(c) and 4(d)]. When H_{ext} is

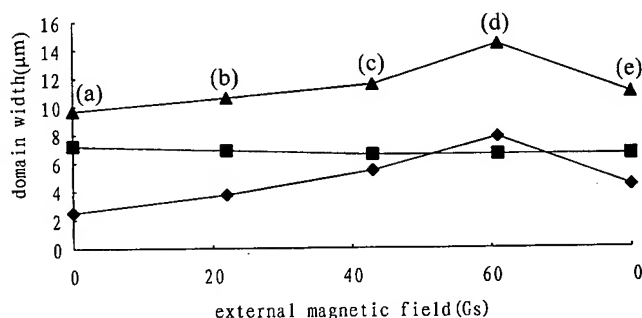


FIG. 5. The relationship of the widths of the domains with the external magnetic field. -◆- bright domain, -■- dark domain, -▲- image contrast periodicity.

raised to 75 G (image not shown), there is no discernable image contrast in the MFM image, which indicates that the whole garnet has been saturated completely.

Because the magnetization direction of hard tip in this case is antiparallel to H_{ext} , the effects they have on magnetizing the garnet is in the cancellation of each other. The effective external field should be, $H_{\text{eff}}=H_{\text{ext}}-H_{\text{tip}}$. Since the tip-sample separation is kept constant during MFM scanning, the impact of the hard tip on the garnet is fixed and tends to reduce the widths of the bright domains. During our static magnetization process, the width of the bright domains is always becoming wider with H_{ext} increasing, indicating that H_{ext} is dominating. On the other hand, the magnetization density of the hard tip could be reduced, which only affects the overall image contrast.

In Fig. 5, the fact that the periodicity of the image contrast changes greatly with increasing H_{ext} suggests that the changes of the domain configuration in Fig. 4 are mainly induced by the external magnetic field.¹² The external field magnetizes the garnet so as to enlarge the widths of the bright domain whose magnetization direction is the same as the external field. Therefore, it also induces changes of the domain configuration.

After the external field is removed, the MFM image of the garnet is shown in Fig. 4(e). Comparing with Fig. 4(a), the domain configuration has not recovered to its original one. This indicates that there are irreversible processes in the course of static magnetization of the garnet.

IV. CONCLUSION

In summary, we have used MFM to observe the process of the variation of magnetic domains in garnet within an external field. The process is observed to undergo the movement of domain walls and the variation of the domain configuration. We have demonstrated that, after removing the external field, the garnet usually does not restore to its original domain configuration, suggesting the existence of irreversible processes in the course of magnetization. We have also studied the effect of magnetic tip on magnetic domains, and illustrated that the effects could be minimized by using a soft magnetic tip and choosing appropriate hard tip-sample distance. The study suggests that the detailed magnetic nature of the tip could serve to further review the insight of the dynamic magnetic processes.

ACKNOWLEDGMENTS

This work was supported by NSFC and the Chinese Academy of Sciences. The authors are indebted to Professor Baoshan Han for providing the sample.

¹P. Grütter, MSA Bull. **24**, 416 (1994).

²D. Rugar, H. J. Mamin, P. Guethner, S. E. Lambert, J. E. Stern, I. McFadyen, and T. Yogi, J. Appl. Phys. **68**, 1169 (1990).

³H. J. Mamin, D. Rugar, J. E. Stern, B. D. Terris, and S. E. Lambert, Appl. Phys. Lett. **53**, 1563 (1988).

⁴C. Schönenberger, S. F. Alvarado, S. E. Lambert, and I. L. Sanders, J. Magn. Magn. Mater. **93**, 123 (1991).

⁵Y. Martin, D. Rugar, and H. K. Wickramasinghe, Appl. Phys. Lett. **52**, 244 (1988).

- ⁶P. Grütter, A. Wadas, E. Meyer, H.-R. Hidber, and H.-J. Güntherodt, *J. Appl. Phys.* **66**, 6001 (1989).
- ⁷C. L. Bai, *et al.*, *J. Phys. Chem.* **99**, 8202 (1995).
- ⁸R. B. Proksch, T. E. Schäffer, B. M. Moskowitz, E. D. Dahlberg, D. A. Bazylinski, and R. B. Frankel, *Appl. Phys. Lett.* **66**, 2582 (1995).
- ⁹R. R. Katti, P. Rice, J. C. Wu, and H. L. Stadler, *IEEE Trans. Magn.* **28**, 2913 (1992).
- ¹⁰A. Wadas, J. Moreland, P. Rice, and R. R. Katti, *Appl. Phys. Lett.* **64**, 1156 (1994).
- ¹¹M. Lederman, G. A. Gibson, and S. Schultz, *J. Appl. Phys.* **73**, 6961 (1993).
- ¹²M. Lederman, D. R. Fredkin, R. O'Barr, S. Schultz, and M. Ozaki, *J. Appl. Phys.* **75**, 6217 (1994).
- ¹³S. Manalis, K. Babcock, J. Massie, and V. Elings, *Appl. Phys. Lett.* **66**, 2585 (1995).
- ¹⁴R. D. Gomez, R. Madabhushi, I. D. Mayergoyz, L. J. Lising, and E. R. Burke, *J. Magn. Magn. Mater.* **148**, 446 (1995).
- ¹⁵S. L. Blank, *J. Cryst. Growth* **17**, 302 (1972).
- ¹⁶W. D. Zhong, *Ferromagnetism* (Scientific, Beijing, 1992) (in Chinese).
- ¹⁷P. Grütter, D. Rugar, H. J. Mamin, G. Castillo, C.-J. Lin, I. R. McFadyen, R. M. Valletta, O. Wolter, T. Bayer, and J. Greschner, *J. Appl. Phys.* **69**, 5883 (1991).

Field induced antiferromagnetic ordering in NiS cluster confined in zeolite Y

Weiping Ding, Wei Zhong, Ning Zhang, Zaibing Guo, Gang Gu, and Youwei Du
*Laboratory of Solid State Microstructures and Department of Physics, Nanjing University,
Nanjing 210093, China*

(Received 12 September 1996; accepted 31 March 1997)

An interesting field induced antiferromagnetic ordering in NiS clusters confined in Y molecular sieve was observed. The NiS cluster-zeolite Y composite exhibits an unusual antiferromagnetic ordering state under a magnetic field of 0.724–1.164 T and paramagnetic behavior out of this field region. The antiferromagnetic ordering can be stable up to 180 K. The phenomenon may be understood by the concept of energy level crossing due to intracluster interaction at a certain magnetic field. © 1997 American Vacuum Society. [S0734-211X(97)04404-1]

I. INTRODUCTION

Recently, much attention has been given to new kinds of materials made by filling the pore space of zeolite with other solid materials, e.g., metals, semiconductors, or polymers.^{1,2} The size quantization effects observed in small clusters in zeolite are currently of great interest for potential nonlinear optic applications. This kind of nanometer structure, metal, or semiconductor constructed in zeolite crystalline, has a lot of researchers excited about zeolites' applications.^{3–5} Already under development are very sensitive chemical sensors and light sensitive materials such as optical memory.⁶ A patent to Ozin and co-workers disclosed a variety of qualitative optical responses including baro-, hydro-, and photo-, and thermochromism for the sodalite family (AlSiO_4)₃Na_{4–n}Ag_nX (with Ag_nX cluster in the pores of sodalite), and reversible read-write cycles with thermal erase were also carried out.⁷ Nozue *et al.* have reported the strong ferromagnetic property exhibited by sodium metal in A-type zeolite.⁸ Additionally, the package of clusters in zeolite crystalline into nanocomposite bulk materials, defining the confinement surface states and environment, intercluster interactions and the quantum-confinement geometry, are also interesting.⁹ Here we report the magnetic field induced antiferromagnetic ordering in NiS cluster confined in Y zeolite. It would be of interest, for understanding various properties of this kind of composite material and some novel applications.

II. EXPERIMENT

The assembly of clusters in zeolite, e.g., II–VI group semiconductor, was usually made by an ion exchange followed by treatment with dry H₂S or H₂Se. It has been observed that moisture has great influence on these samples.^{10,11} However, Barnakov and co-workers have recently reported that the (CdS)-X and (CdS)-A composite material obtained by treating the Cd²⁺ exchanged single crystal X- and A-type molecular sieve with 1 M Na₂S aqueous solution have high stability in air.¹² In consideration of practical utility, our (NiS)-Y sample was prepared by treating the NiY with Na₂S aqueous solution. The NiY samples were prepared by a standard ion-exchange procedure. The starting

material NaY (2g) was slurried into 100 ml of distilled water. The pH was adjusted to 5, and an appropriate amount of nickel nitrate to yield stoichiometries as listed below as added. The slurry was stirred overnight at room temperature, filtered, and washed. The dried sample was heated at a rate of 2 K/min to 573 K for 3 h in air (denoted as NiY1). The color of the NiY1 sample is pale green. The NiY1 was treated in 200 ml 0.5 M Na₂S aqueous solution at 50 °C for 2 h, then the sample was washed, and dried at 50 °C (denoted as NiSY1). After sulfidation, the color of the sample from green (NiY1) turns to gray (NiSY1). Chemical analysis (the Material Center of Nanjing University) reveals the following composition of the unit cell: starting NaY Na(52)Al(52)Si(140); NiY1 Ni(4.6)Na(42)Al(52)Si(140); NiSY1 Ni(3.8)S(3.2)Na(50.9)Al(52)Si(140).

III. RESULTS AND DISCUSSION

The x-ray diffraction patterns showed that the crystallinity of these samples was comparable to that of the parent NaY. The magnetic properties of the samples were measured using a vibrating sample magnetometer (VSM). The powder sample was cooled in zero field down to 77 K and the magnetization was measured as a function of increasing field strength up to 1.5 T. The field was again set to zero and the sample was warmed to a higher temperature. In this manner, measurements were taken as a function of field for a number of temperatures between 77 K and room temperature. Magnetization versus temperature curves were obtained at a fixed magnetic field with temperature programmed to increase. The x-ray photoelectron spectroscopy measurements were performed by using ESCALAB MKII with Mg K α radiation as x-ray source, and the residual carbon C 1s = 284.6 eV was set as reference of binding energy.

The change in the Ni 2p binding energy spectra before and after sulfidation, as shown in Fig. 1, shows that the sulfidation gives rise to the formation of nickel sulfidelike compound.^{13,14}

Moller, Herron, and co-workers have made careful studies on cadmium sulfide and selenide clusters stabilized in zeolite Y by means of EXAFS and high-resolution synchrotron radiation XRD. They found the geometric arrangements of

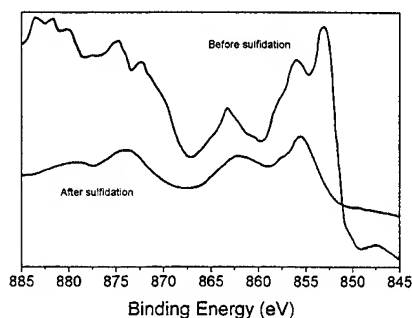


FIG. 1. Ni 2p binding energy spectra before and after sulfidation.

cadmium sulfide in a sodalite subunit of the zeolite framework. The structures can be explained as zeolite stabilized cubelike $(\text{CdS})_4$ with the extra coordination of Cd atoms with the framework oxygen atoms of the six-ring windows. Quantum well tunneling through the aluminosilicate cage wall is also suggested to be responsible for the interaction between clusters.^{10,15-17} Considering the similarity between the chemical property of CdS and NiS, it seems reasonable to speculate that the structure of NiS clusters in Y zeolite is analogous to that of CdS confined in zeolite Y.

Figure 2 gives the variation of magnetization of NiY1 with magnetic field under different temperatures. That the magnetization of NiY1 changes linearly with magnetic field up to 1.5 T shows the paramagnetic behavior of NiY1. It is of great interest to note that, after sulfidation, a dip appears in the $M-H$ curve of NiSY1 (Fig. 3), indicating certain antiferromagnetic ordering induced by a magnetic field. The antiferromagnetic interaction begins at 0.724 T (denoted as H_1), reaches its maximum at 0.936 T (H_0), and vanishes at about 1.164 T (H_2). The samples are stable in air for several weeks and these results are found to be reversible. Beyond the magnetic field region of H_1-H_2 , the NiSY1 displays paramagnetic behavior. With the temperature increase, the dip weakened and disappeared completely at 200 K. The analysis is further supported by the $M-T$ curve measure-

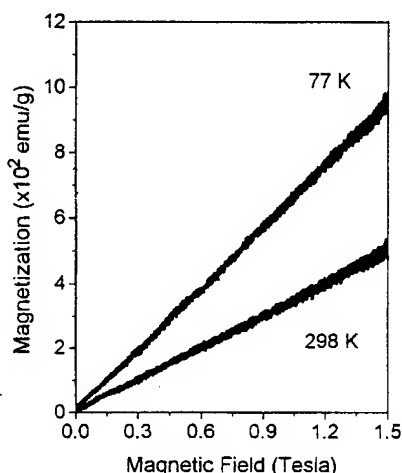


FIG. 2. $M-H$ curve of NiY1 at different temperatures.

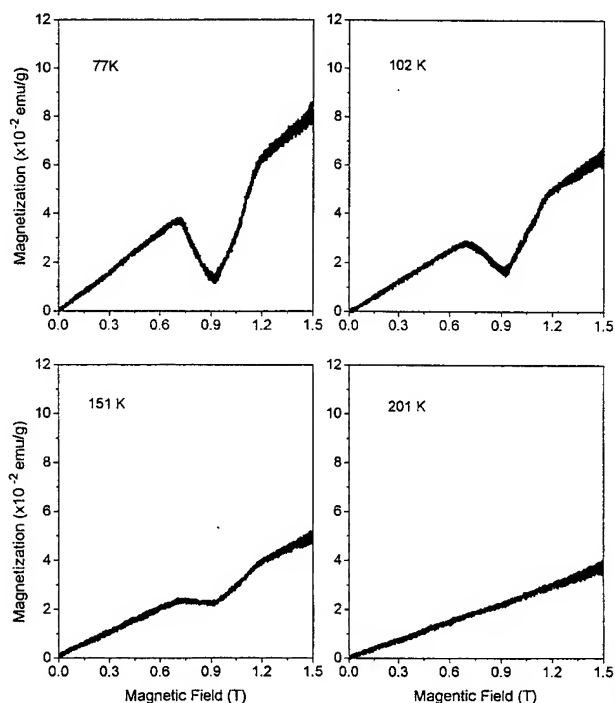


FIG. 3. $M-H$ curve of NiSY1 at different temperature.

ment performed at different magnetic field. As shown in Fig. 4, the variation of magnetization of NiSY1 with temperature has paramagnetic characteristics at a magnetic field of 0.724 T (H_1) and 1.164 T (H_2), respectively. However, NiSY1 sample has antiferromagnetic characteristic at a magnetic field of 0.936 T. The Neel temperature of the antiferromagnetic structure is around 180 K judged by the turning point of $M-T$ curve under 0.936 T magnetic field. It is known that the hexagonal NiS undergoes a first-order metal-nonmetal transition at $T_i \sim 260$ K;¹⁸ the high-temperature phase is a Pauli paramagnetic metal and the low-temperature phase is an antiferromagnetic semiconductor or semimetal. This transition has previously been explained in terms of electron

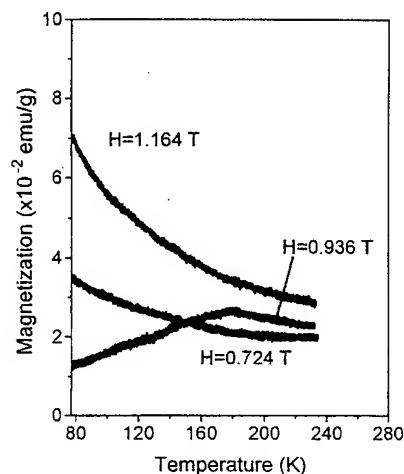


FIG. 4. Variation of magnetization of NiSY1 with temperature at different magnetic fields.

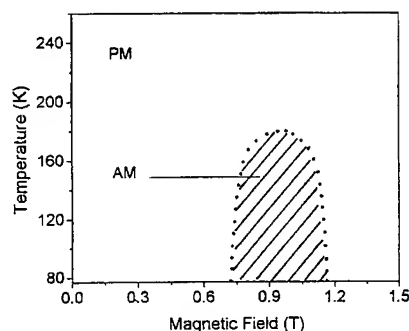


FIG. 5. The phase boundary of NiS clusters confined in zeolite Y determined from $M-H$ curve measurements at different temperatures.

correlation within the Ni 3d band. Namely, the transition has been attributed to the opening of a Mott-Hubbard gap in the nonmetallic phase. Obviously, the experimental facts observed cannot be expounded using that theory. Tsuneto *et al.*¹⁹ have pointed out, through a theoretical analysis, that a magnetic system consisting of ions with spin $S=1$, which have a large uniaxial anisotropy energy D and interact with neighbors by means of an exchange interaction J much smaller than D , shows an interesting spin ordering in an external magnetic field $\mu_B H \sim D$. Ni^{2+} ion ($3d^8$, $S=1$) is one of the non-Kramers ions, which has the large magnetic anisotropy energy in a crystal field. Wada *et al.* have reported an observation of field induced spin ordering in a weakly coupled-pair system: single crystal $\text{Ni}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$ between 39 (H_1) and 49 kOe (H_2) and below 0.1 K.²⁰ By a simple calculation performed on a pair model, i.e., two kinds of Ni^{2+} ions, respectively, form a spin pair with an antiferromagnetic interaction intrapair, they found the ground two levels cross with each other due to intrapair interaction at H_1 and H_2 , respectively. Between the two magnetic fields, the intrapair antiferromagnetic interaction originates the antiferromagnetic ordering. It seems that, qualitatively, a similar concept can be developed to explain the behavior of NiS clusters positioned in the pore of Y zeolite, whereas the NiS clusters confined in Y molecular sieve forms a strong-coupled system considering the antiferromagnetic stability of 180 K. An intracluster antiferromagnetic interaction, like that in bulk NiS, leads to levels crossing under the magnetic field. The magnetic phase boundary of the NiSY1 sample may be roughly depicted in Fig. 5.

It has been recognized that the CdS clusters fabricated in zeolite Y forms supercluster by "percolation effect" at high CdS content. The percolation threshold for CdS in Y mo-

lecular sieve is 4 ± 1 wt %. As this 3D interconnection proceeds, the corresponding changes in optical properties occur.¹⁷ According to the above concept, the NiSY1 sample undergoes two magnetic transitions at H_1 and H_2 , respectively, and shows paramagnetic property at a magnetic field less than H_1 or more than H_2 . It should be interesting to know what happens with increasing the NiS cluster amount confined in Y molecular sieve. Further work is needed.

In conclusion, we have found an interesting field induced magnetic phase transition in NiS clusters confined in Y molecular sieve. To our knowledge, the present system is the first example of strong-coupled, magnetic field induced antiferromagnetic ordering, which indicates again that zeolite-cluster composite can provide a wealth of new materials for which unusual physical properties might be expected. Further studies of the interaction between the clusters confined in the pores of molecular sieve in frame of magnetism should increase our understanding of various properties of this kind of composite material and could offer some novel applications, such as optical, electric, or magnetic devices.

¹G. D. Stucky and J. E. Mac Dougall, *Science* **247**, 669 (1990); S. Feng and T. Bein, *Nature* **368**, 834 (1994); P. A. Anderson and P. Edwards, *J. Am. Chem. Soc.* **114**, 10 608 (1992).

²G. A. Ozin, A. Kuperman, and A. Stein, *Angew. Chem.* **101**, 373 (1989); J. E. Mac Dougall *et al.*, *J. Am. Chem. Soc.* **111**, 8006 (1989).

³G. Schulz-Ekloff, in *Zeolite Chemistry and Catalysis*, edited by P. A. Jacobs (Elsevier, New York, 1991), p. 65.

⁴G. A. Ozin, R. A. Prokopowicz, and S. Özkaz, *J. Am. Chem. Soc.* **114**, 8953 (1992).

⁵V. R. Ramamurthy, J. V. Caspar, D. F. Eaton, E. W. Kuo, and D. R. Corbin, *J. Am. Chem. Soc.* **114**, 3882 (1992).

⁶T. Bein, K. Brown, G. C. Frye, and C. J. Brinker, *J. Am. Chem. Soc.* **111**, 7640 (1989).

⁷G. A. Ozin, J. P. Godber, and A. Stein, US Patent No. 4,942,119 (1990).

⁸Y. Nozue *et al.*, *Stud. Surf. Sci. Catal.* **84**, 2285 (1994).

⁹C.-G. Wu and T. Bein, *Stud. Surf. Sci. Catal.* **84**, 2269 (1994).

¹⁰N. Herron *et al.*, *J. Phys. Chem.* **92**, 4988 (1988).

¹¹M. Wark, G. Schulz-Ekloff, and N. I. Jaeger, *Catal. Today* **8**, 467 (1991); *Stud. Surf. Sci. Catal.* **70**, 189 (1991).

¹²Y. A. Barnakov *et al.*, *Stud. Surf. Sci. Catal.* **84**, 829 (1994).

¹³R. Cid, J. L. G. Fierro, and A. L. Agudo, *Zeolites* **10**, 59 (1990).

¹⁴W. J. J. Welters, G. Vorbeck, H. W. Zandbergen, J. W. Haan, V. H. J. Beer, and R. A. Santen, *J. Catal.* **150**, 155 (1994).

¹⁵Y. Wang and N. Herron, *J. Phys. Chem.* **91**, 257 (1987).

¹⁶N. Herron, Y. Wang, M. M. Eddy, G. D. Stucky, K. Moller, T. Bein, and D. E. Cox, *J. Am. Chem. Soc.* **111**, 530 (1989).

¹⁷K. Moller, M. M. Eddy, G. D. Stucky, N. Herron, and T. Bein, *J. Am. Chem. Soc.* **111**, 2564 (1989).

¹⁸J. Sparks and T. Komoto, *Phys. Lett. A* **25**, 398 (1967); S. Anzai and K. Ozawa, *J. Phys. Soc. Jpn.* **24**, 271 (1968).

¹⁹T. Tsuneto and T. Murao, *Physica* **51**, 186 (1971).

²⁰N. Wada, K. Matsumoto, K. Amaya, and T. Haseda, *J. Phys. Soc. Jpn.* **47**, 1061 (1979).

Study of the complex of the Schiff-base polymer with sulfate iron by magnetic force microscopy*

F. Tian, C. Wang, C. L. Bai, W. G. Li, and M. X. Wan

Institute of Chemistry, The Chinese Academy of Sciences, Beijing 100080, China

(Received 8 September 1996; accepted 2 May 1997)

The micromagnetic distribution of the clusters of the PPH-FeSO₄ complex (Schiff-base polymer with sulfate iron) was investigated by magnetic force microscopy (MFM). MFM results exhibited that magnetites (Fe₃O₄) were irregularly included in PPH-FeSO₄ complexes. Some nanocrystalline Fe₃O₄ magnetites aggregated and formed magnetic clusters of 50–140 nm length. The ferromagnetic properties of the PPH-FeSO₄ complex were associated with the existence of Fe₃O₄ magnetites. Moreover, the influence of the distance between tip and sample surface on the MFM images was also discussed. The observation suggested that MFM was a very useful tool for studying the micromagnetic structure of organic ferromagnets. © 1997 American Vacuum Society.

[S0734-211X(97)13104-3]

I. INTRODUCTION

Magnetic force microscopy (MFM) is a new and useful analytical tool for directly imaging the fine magnetic structure of samples with the advantages of high resolution and no special requirement for sample preparation. Recently, MFM has been extensively applied to study magnetic recording media, magnetic thin films, and some magnetic phenomena in magnetism.^{1,2} However, little study of organic ferromagnets with MFM has been reported.³ Although electron spin resonance (ESR) and superconducting quantum interference device (SQUID) methods have been used to study the macroscopic magnetic properties of organic ferromagnets; these techniques are not sufficient to measure the local magnetic behaviors of the surface of materials.

The complexes of the Schiff-base polymer with sulfate iron (II) PPH-FeSO₄ {[Fe(C₁₃H₁₇N₃)SO₄·6H₂O]_n} have been synthesized and analyzed at our institute.^{4,5} The PPH polymer itself does not exhibit ferromagnetic properties. However, PPH-FeSO₄ displays ferromagnetic properties at room temperature. The coercive force (H_c) and residual magnetization (M_r) for PPH-FeSO₄ complexes at room temperature are 200 Oe and 1.48 emu/g, respectively. The studies by Fourier transform infrared (FTIR), x-ray photoelectron spectroscopy (XPS), x-ray diffraction, Mössbauer, and magnetic measurements confirmed that there exist Fe₃O₄ magnetites in PPH-FeSO₄ complexes.⁶ It is therefore interesting to investigate the fine micromagnetic structures of the PPH-FeSO₄ complex and the distribution of Fe₃O₄ magnetites in the PPH-FeSO₄ complex. In this work, we used MFM to examine the distribution of Fe₃O₄ magnetites in PPH-FeSO₄ complexes directly. Fe₃O₄ magnetites were observed to be embodied in PPH-FeSO₄ complexes and some of them aggregated and formed submicron magnetic clusters. The presence of Fe₃O₄ magnetites resulted in the ferromagnetic behaviors of PPH-FeSO₄ complexes.

II. EXPERIMENT

The PPH-FeSO₄ complexes were synthesized by a method described in Ref. 4. In the PPH-FeSO₄ complex, the ratio of Fe₃O₄ to PPH polymer was 2.714, the iron content was 9.8% and Fe(III) was 4.9% as determined by inductively coupled plasma (ICP) spectra.⁶ IR spectra, XPS, and x-ray diffraction were used for characterization of the structure of the complex.⁶ The elemental analysis suggested that two units of PPH combine with one iron atom to form an octahedral complex. Small amounts of PPH and PPH-FeSO₄ complex powder were dispersed in acetone solution and ultrasonated about 30 min. Later, 20 μl of the mentioned solutions were dropped onto newly cleaved mica. They were then used in the MFM experiment after acetone volatilized completely.

The principle of MFM is quite similar to that of atomic force microscopy (AFM). In MFM, a nonmagnetic Si tip coated with a thin ferromagnetic film was used. With MFM working in its raster scan mode, the tip was brought into contact with the sample surface and performed one scan line first. During this process, the force sensor detected the short-range atomic repulsive force between tip and sample. At the completion of this scan line, the tip was scanned again along a path parallel to the previously stored scan line at a predetermined height from the sample surface. During this second scan the tip-sample separation was nominally constant. At this time, the force sensor detected the long-range magnetic force felt by the magnetic tip owing to stray field of the sample. Therefore, the topography and the magnetic force images of the sample of the same area were obtained as a function of position simultaneously.

If the magnetic interaction between tip and the stray field emanating from the sample was attractive, the MFM image will show reduced contrast, while in the case of repelling interaction, bright contrast appears.

MFM experiments of PPH-FeSO₄ complexes were performed by using a NanoScope III a Dimension™ 3000 scanning probe microscope, manufactured by Digital Instruments Inc. The cantilever was a commercial NanoProbe magnetically coated tip with a length of 225 μm and had a resonant

*No proof corrections received from author prior to publication.

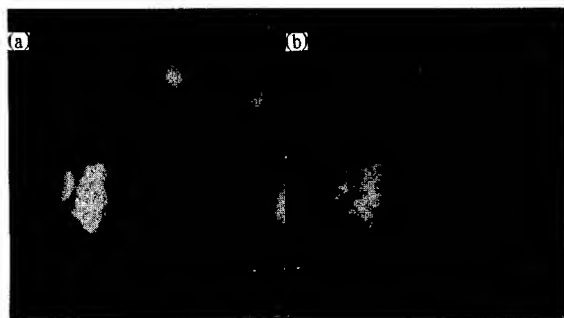


FIG. 1. (a) AFM topography of the PPH polymer. (b) Corresponding image of the polymer taken simultaneously with the AFM topography using a magnetic tip. The tip-sample separation was 200 nm. Scan sizes $6.0 \mu\text{m} \times 6.0 \mu\text{m}$.

frequency of approximately 70 KHz. The tip was magnetized in a magnetizer along the tip direction, which was perpendicular to the imaging surface of the sample. The topographical and magnetic images were obtained simultaneously using tapping and lift modes, respectively.

III. RESULTS AND DISCUSSION

Figure 1(a) shows an AFM topography of PPH polymer particles. Figure 1(b) is the corresponding image of the PPH polymer taken with tip-sample separation of 200 nm using a magnetized tip. It is evident that in Fig. 1(b) the regions with bright contrast closely resemble the shapes of the PPH particles in AFM topography. Figure 2(a) gives an AFM topography of the clusters of PPH-FeSO₄ complexes. Figures 2(b)–2(d) are the corresponding MFM magnetic force images of PPH-FeSO₄ complexes. The tip-sample separations for Figs. 2(b)–2(e) are 200, 100, 50, and 30 nm, respectively. In Fig. 2(a) the ultrasonated PPH-FeSO₄ complex clusters are about 400–800 nm in length and 60–140 nm in height. It is interesting to notice that in Figs. 2(b)–2(e) there are many dark regions whose positions are inside PPH-FeSO₄ complex clusters corresponding to Fig. 2(a). Previously, studies

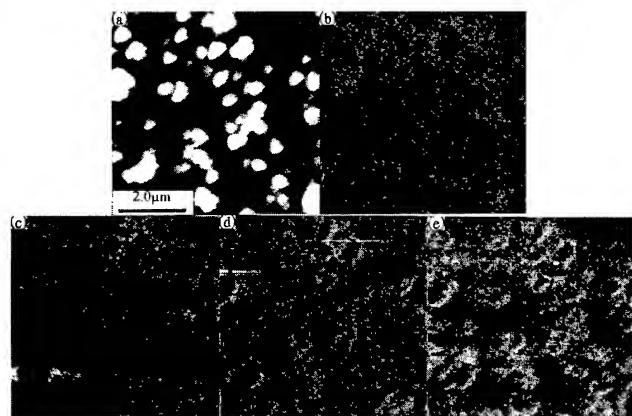


FIG. 2. (a) AFM topography of the clusters of PPH-FeSO₄ complexes. (b)–(e) Corresponding MFM images of PPH-FeSO₄ complexes. The tip-sample separation was (b) 200 nm, (c) 100 nm, (d) 50 nm, (e) 30 nm. The imaged regions were $6.0 \mu\text{m} \times 6.0 \mu\text{m}$.

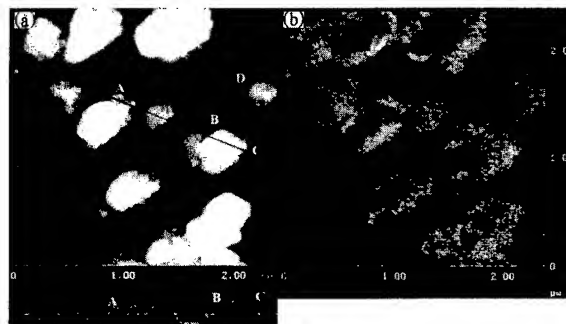


FIG. 3. Magnified (a) AFM and (b) MFM images of PPH-FeSO₄ complexes. The tip-sample separation 30 nm, scan size $2.37 \mu\text{m} \times 2.37 \mu\text{m}$. The cross-sectional profile was taken across the clusters of PPH-FeSO₄ complexes as shown in the figure.

have demonstrated that the PPH polymer itself does not exhibit ferromagnetic property, but, the PPH-FeSO₄ complex displays ferromagnetic properties at room temperature.⁴ The dark regions in Figs. 2(b)–2(e) should be associated with Fe₃O₄ magnetites and some of the bright regions in Fig. 1(b) could be the topographical characteristics. Moreover, in Fig. 1(b), it is clear that the bright regions of the large particles appear at a tip-sample separation as large as 200 nm, while small particles are not. This suggests that the large particles tend to give rise to a topographic feature.

On imaging particles with MFM, the cause for the sample topographical feature appearing in MFM images could be attributed to the presence of nonmagnetic force. As the magnetic tip scans the sample surface, the long-range nonmagnetic interaction force, such as van der Waals force, is added to magnetic force. Therefore, the MFM images are jointly affected by competing force (magnetic and nonmagnetic forces). Since their dependence on separation is very different, it is possible to have a suitable working distance for MFM operations. If the magnetic tip is very close to the sample surface, the magnetic force could be less dominating and the presence of nonmagnetic force makes the characteristics of sample topography more pronounced in MFM image. Comparing Figs. 2(b)–2(e), as the tip-sample distance decreased, the resolution and contrast of dark magnetic regions of magnetite Fe₃O₄ in MFM image enhance. One large blurred area becomes a few small dark magnetic areas. Meanwhile, the topographical feature corresponding to some bright regions around dark regions gradually appear. The reason is that, as tip-sample distance increased, the magnetic force becomes dominating and the influence of nonmagnetic force is minimized. However, the increasing of separation between tip and sample causes the reduction of the magnitude of magnetic force between them, resulting in the reduced contrast in MFM images.

Figure 3 is a small scan of the PPH-FeSO₄ complex. Comparing AFM image [Fig. 3(a)] with MFM image [Fig. 3(b)], one finds that Fe₃O₄ magnetites are irregularly embedded in PPH-FeSO₄ complex clusters, and larger PPH-FeSO₄ complex clusters tend to contain more Fe₃O₄ magnetites while less or no Fe₃O₄ magnetites exist in small

PPH-FeSO₄ complex clusters. As a result, these Fe₃O₄ magnetites form dark magnetic clusters which are about 50–140 nm in length. The dimension is relatively large compared with the previously reported result⁵ in which the size of nanocrystalline Fe₃O₄ magnetites was reported to be 15–20 nm in diameter. Taking into account the resolution of MFM and the influence of tip geometry, we propose that some nanocrystalline of Fe₃O₄ magnetites in the clusters of the PPH-FeSO₄ complex might aggregate and form small magnetic clusters, and the tip might not discern individual particles from each other in this study.

It should be noted that different orientation of magnetization may result in drastic difference in MFM image contrast.⁷ Therefore, some of the bright regions in MFM images of PPH-FeSO₄ complexes may be caused by aggregated Fe₃O₄ magnetites, too. By comparing the height variations, one could reasonably identify the possible locations of magnetites. As an example, positions marked as A, B, C, D in the cross-sectional profile of [Fig. 3(a)] display different contrast in the corresponding MFM image [Fig. 3(b)]. Therefore, position A is believed to be associated with Fe₃O₄ magnetite while position D is only a topographic feature.

IV. CONCLUSION

The fine distribution of Fe₃O₄ magnetites in the clusters of PPH-FeSO₄ complexes have been studied by using MFM. The Fe₃O₄ magnetites are observed to be irregularly included in PPH-FeSO₄ complexes and nanocrystalline Fe₃O₄ magnetites tend to an aggregate and form magnetic clusters of 50–140 nm length in dimension.

During MFM experiments, it is very important and necessary to have good understanding and control of reasonable tip-sample operation distances in order to minimize non-magnetic contributions to the tip-sample interaction and make a correct interpretation to MFM experimental data. In any case, MFM is proven to be a very useful tool to study the micromagnetic structure and the distribution of magnetites in organoferromagnets. If we can verify the nature of the magnetic tip, the magnetic field from PPH-FeSO₄ complexes could be quantitatively studied. The observation of this letter serves to lay a ground for further investigation of the magnetic nature of organoferromagnets.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China and the Foundation of The Chinese Academy of Sciences.

¹C. L. Bai, *Scanning Tunneling Microscopy and Its Application* (Springer, Berlin, 1995), p. 116.

²P. Grütter, *MSA Bull.* **24**, 416 (1994).

³C. Bai, P. Zhang, D. Zhu, M. Han, Y. Xu, D. Zhang, and Y. Liu, *J. Phys. Chem.* **99**, 8202 (1995).

⁴M. X. Wan, Ch. H. Zhang, and Ch. J. Li, *Solid State Commun.* **87**, 379 (1993).

⁵M. Wan, J. Li, P. Auric, G. Legaer, B. Malaman, and E. Ressouche, *Solid State Commun.* **89**, 999 (1994).

⁶W. G. Li and M. X. Wan, *Solid State Commun.* **92**, 629 (1994).

⁷F. Tian, C. Wang, G. Y. Shang, N. X. Wang, and C. L. Bai, *J. Magn. Magn. Mater.* (in press).

Imaging, polymerization, and reconstruction of polystyrene films with a scanning tunneling microscope

Z. Y. Hua^{a)} and W. Xu

Department of Materials Science, Fudan University, Shanghai 200433, China

(Received 6 March 1997; accepted 10 March 1997)

Polystyrene microparticles can be investigated by using a scanning tunneling microscope. The observed images show the hydrogen atoms on the phenyl groups and the extended polymer chains without entanglement. After a sustained tip scanning, some of the polystyrene surfaces develop a highly ordered structure. Furthermore, the polymerization of styrene monomers can be initiated by the tip field, to form nanometer-scale polystyrene films on graphite substrates, in air at room temperature. The yield was verified by time-of-flight secondary ion-mass spectroscopy. © 1997 American Vacuum Society. [S0734-211X(97)01504-7]

I. INTRODUCTION

A scanning tunneling microscope (STM)¹ can give atomic resolution imaging of solid surfaces which have, in principle, sufficient electrical conductivity. Otherwise an atomic force microscope (AFM)² is preferable. However, in recent years, it has been found that some very thin polymer films, e.g., polyethylene,^{3,4} poly(1-butene),⁵ polybutadiene,⁶ poly(ethylene oxide),⁷ and single-molecule, single crystal of isotactic polystyrene⁸ can be imaged by STM. For poly(methyl methacrylate), polydiphenyl-siloxane, and conventional polystyrene,⁹ surface morphology can be shown without molecular details. There is still a lack of clear understanding of the mechanism, however, in the case of polystyrene, the image resolution apparently depends upon physical and chemical states of the polymers. This article reports our studies on surfaces of polystyrene microparticles (PM). Their structures can be imaged with high resolution and reconstructed under sustained operation of a STM.

Besides the imaging function, the electric field of the STM tip under normal operating conditions has drawn our attention. The tip-sample distance is only on the order of 0.1 nm, therefore, even though the voltage applied between them is very low (<1 V), the electrostatic field produced is extremely high. This field has already induced the tunneling current for STM imaging,¹⁰ and proved to be a good means for submicron processing and/or atom manipulation.¹¹ This article examines some additional effects, e.g., surface polymerization and surface reconstruction. It is remarkable that a high-resolution, real-time observation is possible during the evolution of a surface. This is unique and thus makes the STM a multifunctional tool.

II. IMAGING OF POLYSTYRENE MICROPARTICLE SURFACES

The conventional atactic polystyrene with entangled polymer chains has never been seen by atom-resolved STM, yet the single molecule, single crystal of isotactic polystyrene can give a rough STM image. Very recently we have applied

STM to the investigation of surfaces of PM prepared by microemulsion free-radical polymerization,¹² where each particle contains only one or a few high molecular weight polymer chains which are strictly constrained in a very limited space (about 20 nm in diameter). The PM has been found to have different properties from ordinary atactic polystyrene.¹³⁻¹⁵ In our case, the molecular weight of the PM is about 1×10^6 .¹⁵

Highly oriented pyrolytic graphite (HOPG) was selected as the substrate because it is clean, inert, conductive, and defect-free in a relatively large area. Moreover, since the surface structure of HOPG is universally known, the STM image obtained cannot be confused with the substrate. After mixing PM with de-ionized water by ultrasonic stirring, a drop of the aqueous suspension was deposited on a substrate of freshly cleaved HOPG, then dried naturally at room temperature for a week. A film of PM was formed on HOPG. Scanning electron microscopy (SEM) showed that film surfaces, although rugged, were fully covered with PM.

Our STM is a combination of hardware CSTM-9000 made by the Institute of Chemistry, CAS, and a software written by ourselves. The software has the ability to permit fast data acquisition, fast image processing, and allows four different images to be displayed simultaneously for an effective comparison of surface structural changes. The STM tips were made of Pt-Ir. All images were obtained in air under ambient conditions. Experimental results can be summarized as follows.

(1) The PM surface had different modes of coagulation. In most regions of the surface, the tip current was not stable and the background noise level was very high, hence no STM image could be obtained. However, at least in some regions, the surface could be imaged clearly. Figures 1(a) and 1(b) were two images obtained from a region with different magnification. The STM images sometimes varied from region to region, however, no entangled polymer chains were ever observed.

(2) In high-resolution STM images, generally there are two or three bright spots assembled together to construct groups of spots [Fig. 2(a)]. In a typical group of two spots, the measured distance between them is about 2.5 Å. In a

^{a)} Author to whom correspondence should be addressed.

group of three spots, usually the middle spot is the brightest one, and the maximal distance between two terminal spots can reach 4.3 Å. We proposed that these bright spots might be the STM images of hydrogen atoms on the upward rim of phenyl groups. The phenyls are sitting vertically on the polystyrene surface, hence the hydrogen atoms are higher than other atoms and thus give brighter spots.^{16,17} This statement was further substantiated as follows. Assuming that the structure of a phenyl group is nearly identical to that of a benzene molecule,¹⁸ where the C–C distance is 1.395 Å and C–H distance is 1.08 Å, we can calculate the horizontal distances between the upward H atoms for three different modes of phenyl groups sitting vertically on the surface, as shown in Fig. 2(b). The computed results, d_A , the maximal distance between two terminal spots of a three-spot group, and d_C , the distance between two bright spots of a two-spot group, are 4.30 and 2.48 Å, respectively. When the phenyl group is not sitting symmetrically to the normal of the surface, the distance d_B should be shorter than d_A but longer than d_C . In this case usually one spot is very faint. Therefore, we only measured d_A and d_C as the criterion of H sites. The results are 4.3 and 2.5 Å, respectively, which is in good agreement with the computed values.

(3) Since the phenyl groups are sitting vertically, we cannot see their cyclic structures directly. However, the alignment of the spot groups gives a definite orientation, from which the projection of the phenyl groups on the surface, and

then the orientation of the phenyl plane, can be decided. The ability to determine the orientation of phenyl plane promises great importance in studying the interaction between the phenyl groups.

III. STACKING OF PHENYL GROUPS

Noncovalent interaction between aromatic units plays a major role in determining the properties and the behavior of molecules. Mitchel and Windle¹⁹ suggested the existence of the phenyl stacks in polystyrene glasses according to the measurement of the wide-angle X-ray scattering, yet it has never been seen. Here a direct observation of the stacking has been realized for the first time.

Figure 3(a) shows that on the PM surface many large, elongated spots are arranged in well defined chains. As mentioned previously, these large spots which consist of two or three smaller ones correspond to phenyl groups that are vertically sitting in parallel. This arrangement can be compared to a stack of coins standing on their rims and leaning against one another, in which each coin represents a phenyl group. We believe we have seen the stacking of phenyl groups.¹⁹ It is noticeable that the phenyls stacked with a direction parallel to the axis of the polymer chains; however, this direction is not perpendicular to the phenyl plane but at an angle of about 65°, as shown in Fig. 3(b). This is similar to the proposed "most stable" parallel-displaced (PD) structure^{20,21} for benzene dimers by *ab initio* calculations. In the case of stacking, since more than two phenyls are involved, the stacks appear as the multiple PD structure. Because each

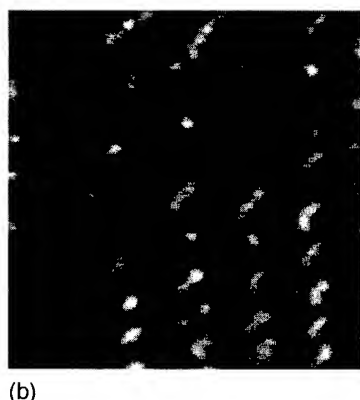
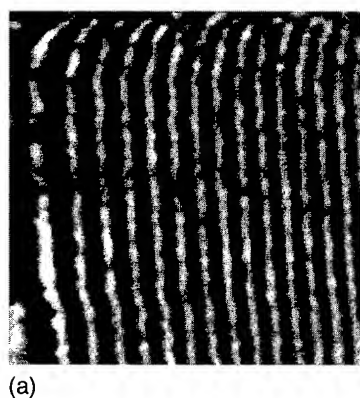


FIG. 1. STM images of a polystyrene surface. (a) 10.5 nm×10.5 nm; (b) 4.0 nm×4.0 nm.

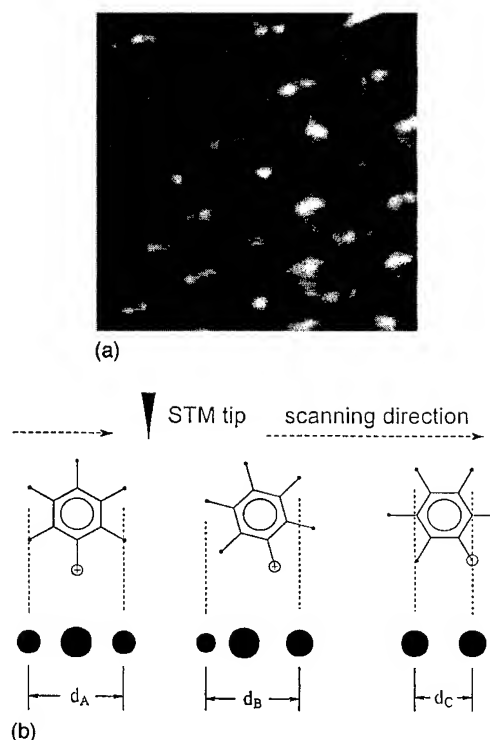


FIG. 2. (a) High resolution STM image of polystyrene surface, 3.7 nm×3.7 nm. (b) The models for explanation of three typical images of hydrogen atom groups. In this schematic drawing, a bigger spot means brighter.

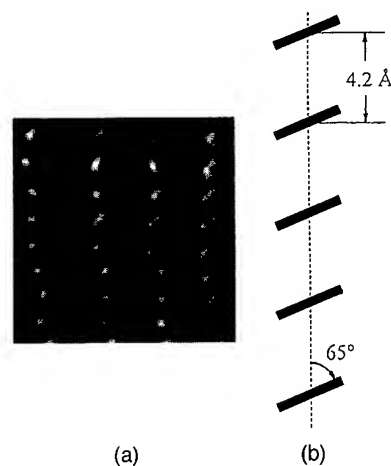


FIG. 3. (a) An STM image of phenyl stacks, 3.7 nm \times 3.7 nm. (b) A model of the multiple PD structure for the stacking of phenyls.

phenyl group participates in two PD structures and the phenyls are bound on the polymer chain, the distance between two neighboring phenyls (4.2 Å) is slightly larger than that in a benzene dimer (3.9 Å).²¹ This kind of stacking may greatly affect the conformation and properties of the polymer.

IV. WAGGING MOTION OF POLYMER CHAINS AND RECONSTRUCTION

In our STM, the information obtained from two directions of tip scanning, i.e., from left to right (L-R) and from right to left (R-L), can be separately collected, and two images may be plotted. In the case of PM, it is found that the STM images on the same surface area dramatically depend on different tip scanning directions, L-R and R-L, as shown in Figs. 4(a) and 4(b), respectively. It implies that the segments of polymer chains can be dragged by the STM tip at the scanning direction, and squeezed to the end terminal of every scan. Obviously they are essentially symmetrical but not identical because the chains are not inherently perpendicular to the scanning direction of the STM tip.

The dragging effect has been ignored when metals, alloys, and oxides are imaged, because the STM tip cannot drag these atoms or molecules under ordinary operating conditions of a STM. For instance, when a bare HOPG is imaged, no dragging effect can be seen. The dragging effect for polymers is not surprising. The effect of drag of the AFM tip has been reported and even molecules can be pulled out by the atomic force.²² Since STM has a very intense electric field, we proposed that the effect may be more dramatic. In the case of polymers like polystyrene, the polymer chains, especially the phenyl groups, are highly polarizable and thus can be easily polarized under the tip field. Consequently, an electrostatic attractive force should exist between the STM tip and the chains. While the tip is moving to-and-fro, the polymer chains are dragged accordingly to cause a wagging motion.

As a direct result of the wagging, under some circumstances the scanning electric field can induce the ordering

surface reconstruction. Figure 5 showed that the surface structure can be varied with time under a sustained tip scanning. At the very beginning [Fig. 5(a)], on the first received STM image, only small spots can be seen. After scanning 2 min, the image [Fig. 5(b)] showed that instead of small ones, numerous larger and longer bright spots were observed. As mentioned previously, these might represent the phenyl groups. To continue the scan makes the images clearer and more orderly. Finally, after 30–40 min, a stable reconstructed surface was available, as shown in Fig. 5(c). The created surface has a closely packed structure. The crystal-like surface suggests the existence of phenyl stacking. It implies that the phenyl groups are not only on the same polymer chain but also on the neighboring chains stack. The arrangements of phenyls in the stacks tends to interlock neighboring chains and accordingly results in the crystal-like structure.

Generally speaking, the successive wagging motion of chains caused by the to-and-fro scanning field gives an opportunity for every phenyl group to arrange according to minimum potential energy requirements, hence it gives the periodic structure. This result also implies that an intense scanning electric field can be used for surface modification.

V. THE INLAYLIKE SELF-ASSEMBLY OF POLYMER CHAINS

We have also investigated the PM swelled with a small amount of benzene. The solution that contained nonsoluble

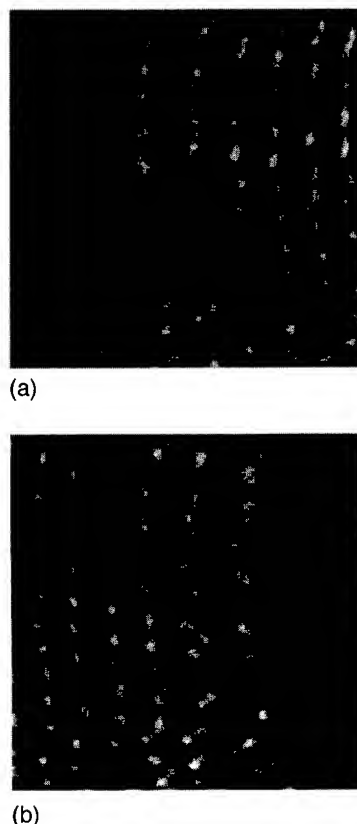


FIG. 4. STM images from the same region of polystyrene surface but with different scanning directions, 5.3 nm \times 5.3 nm: (a) L-R; (b) R-L.

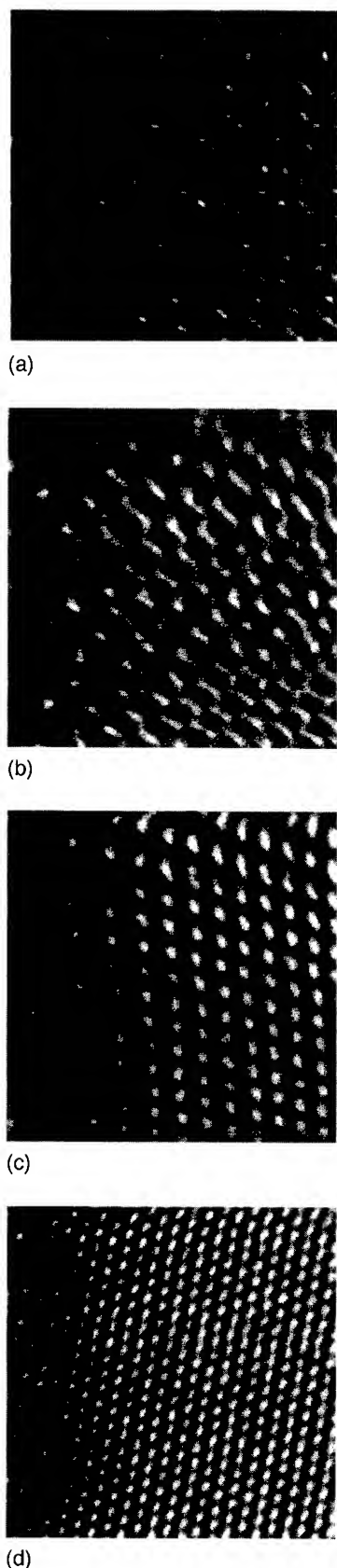


FIG. 5. STM images of the surface reconstruction, $5.3 \text{ nm} \times 5.3 \text{ nm}$: (a) the first received; (b) 2 min later, (c) the final, 0.5 h later. An HOPG surface structure (d) of the same size under the same operating conditions, is used for comparison.

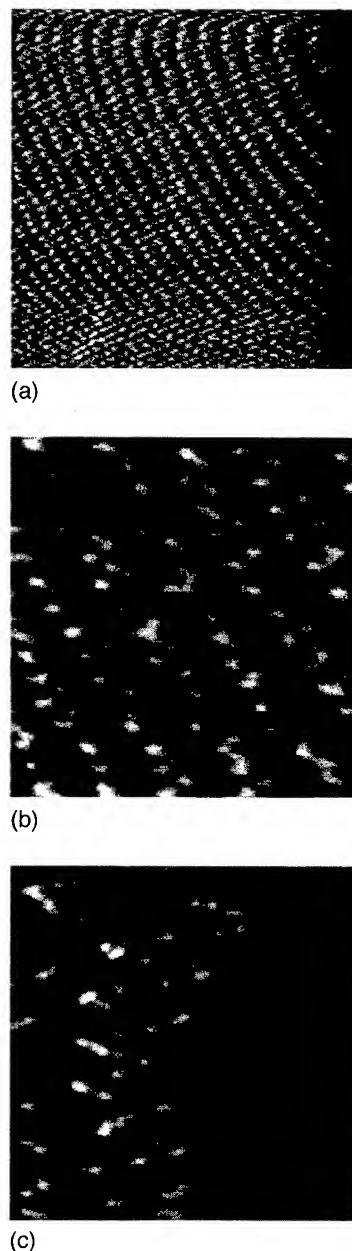


FIG. 6. STM images of swelled polystyrene surface: (a) $19.3 \text{ nm} \times 19.3 \text{ nm}$, (b) $7.5 \text{ nm} \times 7.5 \text{ nm}$, (c) $5.3 \text{ nm} \times 5.3 \text{ nm}$.

particles was deposited on a HOPG surface. After drying, the STM images also revealed different modes of surface coagulation. An image with a relatively larger scale is presented in Fig. 6(a). The bright spots arranged as a chainlike structure might be the extended polymer chains. Further decreasing the scanning area for a better resolution, the fine structure can be seen, as shown in Fig. 6(b), where the polymer chains self-assembled as an inlaylike form. The image with the best resolution is shown in Fig. 6(c), where the inlaylike self-assembly seems more distinct. But, since the chains are dragged by the STM tip and squeezed onto one side of the image, only two chains can be seen in this scanned area.

From Figs. 6(b) and 6(c), we can see that the shape of polymer chains is similar to helixes. The measured pitch of

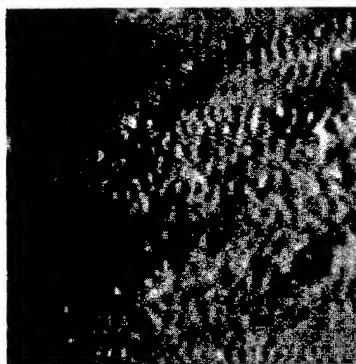


FIG. 7. STM image of the region which was directly under the intense electric field, $10.5 \text{ nm} \times 10.5 \text{ nm}$.

the helix is about 0.6 nm , which is in good agreement with the measured value for c axis in a crystalline cell of isotactic polystyrene by x-ray diffraction.²³ This is interesting, but no satisfactory explanation is available now.

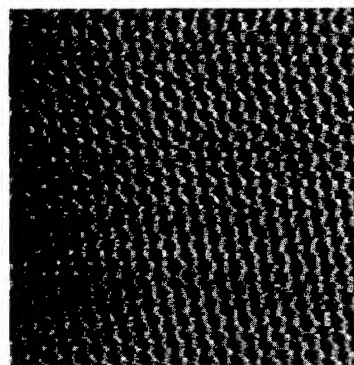
VI. FIELD-INITIATED POLYMERIZATION OF STYRENE

Considering that the STM tip can provide an extremely intense electrostatic field (10^9 – 10^{10} V/m), it may be used for some other purposes. Under optimal conditions, no matter how strong this field, the energy of tunneling electrons is relatively low (0.1 – 1 eV), by which no effect of electron bombardment should be considered. All the effects which happened, if any, should be due to the field itself. Therefore, this is a very good condition to verify pure field effects, and STM itself might be the unique instrument to provide microscopic inspection of these effects. The purpose of this work is to investigate whether any reaction will happen to some monomers under such a field.

A thin film was deposited by one drop of distilled styrene (liquid) on an HOPG substrate. The styrene began to evaporate instantaneously, then a very thin film was formed on HOPG. The sample was put into the STM sample chamber in air at room temperature. No initiator or catalyst was added. With a bias of $+0.1$ to $+1.0 \text{ V}$, a Pt/Ir tip was arranged to be close to the sample for operation with conventional constant-current mode. At the very beginning, the tip current was unstable, and the background noise level was very high. Nothing could be observed. But after a couple of minutes, the tunneling current became stable. The STM gradually presented a rugged topographical image, as shown in Fig. 7. When the STM tip was moved to a new neighboring region, the first received STM image showed some flat and uniform surfaces with slightly ordered structures, as shown in Fig. 8(a). When the tip was continuously scanning, the figure became more and more orderly, and finally the surface presented a crystal-like structure, as shown in Fig. 8(b). It implies that the scanned area behaves as a "crystal seed," and the crystalline structure can be automatically self-extended under suitable conditions.



(a)



(b)

FIG. 8. STM images of a reconstructing polystyrene film, $5.3 \text{ nm} \times 5.3 \text{ nm}$: (a) initial image, (b) final image (after sustained scanning for 6 min).

We further discovered that the initiation of polymerization was not dependent on scanning at all. If the tip stayed at any fixed point for 2–5 min with an applied bias voltage, the result was just the same. For confirmation of the product, it was analyzed by a time-of-flight type secondary-ion mass spectrometer (TOF-SIMS). The instrument used in our investigation is a TFS-2000 MPI produced by Charles–Evans and Associates Co. A gallium liquid metal ion source delivers a Ga^+ ion current of 60 pA at 15 keV . Compared with the existing data of static SIMS,²⁴ we found that our product shows all the known characteristic fragments of polystyrene, i.e., negative ion peaks with mass number 37 (C_3H^-), 49 (C_4H^-), 62 (C_5H^-), 73 (C_6H^-); positive-ion peaks of 51 (C_4H_3^+), 63 (C_5H_3^+), 77 (C_6H_5^+), 91 (C_7H_7^+), and relatively weak but important signals of 107, 115, 128, 152, 165, 178, 193, 207, 221, and 281. After analysis the sample was kept in vacuum environment under room temperature for 23 h, and then re-analyzed by TOF-SIMS. The resulting mass peaks were very reproducible.

VII. CONCLUSION

Structures of PM prepared by the microemulsion method differ from conventional atactic polystyrene. The surfaces of PM can be imaged by STM with high resolution. The STM images revealed the existence of the orderly stacked phenyls and polymer chains which is impossible for ordinary polystyrene. The extended polymer chains can be dragged by the

STM tip. Furthermore, the intense scanning field can induce the ordered surface reconstruction. The hydrogen atoms on the phenyl groups are distinguished by atomic resolution imaging, and accordingly the orientation of the phenyl plane can be determined. In some cases, the phenyl groups are found to be stacking as the parallel-displaced (PD) structure.^{20,21} This experimental observation supports the results of *ab initio* calculations for benzene dimers. The inlay-like self-assembly of polystyrene chains is of value in studying the aggregation of polymers. Together with other experimental results, our observation may provide a deeper insight into the PM microstructures.

It was proved that styrene monomers can be directly polymerized to form a thin polystyrene film under an intense electric field, in air at room temperature. Disregarding tip scanning, the active center can be an area or simply a point. However, only after a sustained scanning of field will the polystyrene film on a single-crystal surface have a highly ordered reconstructed structure. This process can be regarded as the growth of epitaxial polymer film on a crystalline substrate.^{25,26} Our experiments show that STM can be a unique versatile instrument to initiate polymerization, provide microscopic imaging, and induce surface reconstruction.

ACKNOWLEDGMENTS

The authors thank R. Shen, T. Y. Yu, and Y. M. Lu for helpful discussions. They also appreciate X. L. Lu for providing the PM sample and Y. F. Chen for his outstanding work on STM instrumentation. This work is partially supported by the Chinese Natural Science Foundation (CNSF) and the Shanghai Science Foundation.

¹G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, *Appl. Phys. Lett.* **40**, 178 (1982).

- ²G. Binnig, C. F. Quate, and Ch. Gerber, *Phys. Rev. Lett.* **56**, 930 (1986).
- ³D. H. Reneker, R. Patil, and S. J. Kim, *Polym. Prepr. Am. Chem. Soc. Div. Polym. Chem.* **33**, 790 (1992).
- ⁴Z. Q. Xue, H. J. Gao, W. M. Liu, C. Zhu, Z. Ma, and S. Pang, *Appl. Surf. Sci.* **60-61**, 346 (1992).
- ⁵M. Fuchs, L. M. Eng, R. Sander, J. Petermann, K. D. Jandt, and T. Hoffmann, *Polym. Bull.* **26**, 95 (1991).
- ⁶W. G. Morris, D. M. White, J. L. Gordon, and T. Thundat, *J. Vac. Sci. Technol. A* **10**, 623 (1992).
- ⁷R. Yang, X. R. Yang, D. F. Evans, W. A. Hendrickson, and J. Baker, *J. Phys. Chem.* **94**, 6123 (1990).
- ⁸H. S. Bu, E. Q. Chen, J. E. Yao, S. Y. Xu, and Y. L. Kuang, *Polym. Eng. Sci.* **32**, 1209 (1992).
- ⁹J. R. Sheats, *Langmuir* **10**, 2044 (1994).
- ¹⁰J. K. H. Horber, W. Haberle, P. Ruppersberg, M. Niksch, D. P. E. Smith, and G. Binnig, *J. Vac. Sci. Technol. B* **12**, 2243 (1994).
- ¹¹L. J. Whitman, J. A. Strosio, R. A. Dragoset, and R. J. Celotta, *Science* **251**, 1206 (1991).
- ¹²J. S. Guo, M. S. El-Aasser, and J. W. Vanderhoff, *J. Polym. Sci., Polym. Chem. Ed.* **27**, 691 (1989).
- ¹³R. Y. Qian, L. H. Wu, D. Y. Shen, D. H. Napper, R. A. Mann, and D. F. Sangster, *Macromolecules* **26**, 2950 (1993).
- ¹⁴Q. Xue, Y. Lu, G. Q. Shi, and Q. P. Dai, *Polymer* **35**, 892 (1994).
- ¹⁵J. Zhao, W. H. Ming, X. L. Lu, C. C. Wang, and S. K. Fu, *Chem. J. Chin. Universities* **16**, 1960 (1995).
- ¹⁶D. P. E. Smith, H. Horber, Ch. Gerber, and G. Binnig, *Science* **245**, 43 (1989).
- ¹⁷D. P. E. Smith, J. K. H. Horber, G. Binnig, and H. Nejo, *Nature (London)* **344**, 641 (1990).
- ¹⁸G. M. Loudon, *Organic Chemistry* (Addison-Wesley, New York, 1984), p. 568.
- ¹⁹G. R. Mitchell and A. H. Windle, *Polymer* **25**, 906 (1984).
- ²⁰P. Hobza, H. L. Selzle, and E. W. Schlag, *J. Phys. Chem.* **97**, 3937 (1993).
- ²¹P. Hobza, H. L. Selzle, and E. W. Schlag, *J. Am. Chem. Soc.* **116**, 3500 (1994).
- ²²O. M. Leung and M. C. Goh, *Science* **255**, 64 (1992).
- ²³G. Natta, *Makromol. Chem.* **35**, 94 (1960).
- ²⁴D. Briggs, A. Brown, and J. C. Vickerman, *Handbook of Static Secondary Ion Mass Spectrometry (SIMS)* (Wiley, New York, 1989).
- ²⁵M. Sano, D. Y. Sasaki, and T. Kunitake, *Macromolecules* **25**, 6961 (1992).
- ²⁶M. Sano, D. Y. Sasaki, and T. Kunitake, *Science* **258**, 441 (1992).

Application of the scanning force microscope in structuring and in temperature-dependent analysis of Au nanostructures

H. Göbel, L. Jacobs, and P. von Blanckenhagen^{a)}
*Forschungszentrum Karlsruhe, Institut für Materialforschung I, Postfach 3640,
D-76021 Karlsruhe, Germany*

(Received 8 September 1996; accepted 12 May 1997)

The amplitudes of one-dimensional periodic profiles were analyzed as functions of annealing temperatures to determine surface diffusion constants and activation energies with high lateral resolution. The temperature dependence of the surface self-diffusion constant was derived for temperatures between 573 and 1073 K by Fourier analysis of the profiles. A temperature-dependent study of the surface of nanocrystalline Au revealed an increase in grain size with rising temperature and the ratio of free energy on the surface and at grain boundaries. © 1997 American Vacuum Society. [S0734-211X(97)13204-8]

I. INTRODUCTION

In the past few years, scanning probe microscopy has achieved crucial importance in materials science in the quantitative determination of surface properties, such as roughness, friction, adhesion, surface diffusion, and grain boundary structure. In addition, surface modification, e.g., structuring and materials deposition, has become an important tool in basic micro- and nanotechnology research.¹

In this study, scanning force microscopy (SFM) is shown to be a versatile tool for modifying metal surfaces in the nanometer range, and for analyzing the thermal stability of nanostructures. A scanning force microscope equipped with a cantilever with a high spring constant has been used in structuring surfaces of bulk nano-, poly-, and single-crystalline gold as well as of ultrathin gold films deposited on various substrates. This report only deals with structuring the surface of polycrystalline gold. The subsequent analysis was performed in the contact mode with a cantilever with a low spring constant. Local measurement of surface self-diffusion constants requires scratches or one-dimensional periodic grids to be produced and their decay by thermally activated diffusion processes to be analyzed. Both procedures can be performed very accurately by means of a scanning force microscope.^{2,3}

The dependence on time and temperature of "single-scratch" surface profiles was studied before at a polycrystalline gold surface.² The data were analyzed on the basis of the theory developed by Mullins.⁴ It is obvious, from that theory, that one-dimensional periodic grids with a small lattice constant lend themselves much better to surface diffusion studies than single scratches, as the profile decay velocity is higher, and the average profile can be determined by Fourier analysis. SFM has been used not only to generate and analyze nanostructures, but also to measure grain sizes on the surface of nanocrystalline gold, and the ratio of free energy on the surface and at grain boundaries by determining the grain boundary surface angle.⁴

II. EXPERIMENTAL DETAILS

A commercial SFM was used in this study (Nanoscope III). Stiff Si cantilevers (spring constant $C_c = 22-79$ N/m) were used for structuring, and the analysis was performed with soft Si_3N_4 cantilevers ($C_c \approx 0.06$ N/m) in the contact mode.

Surface structuring in the nanometer range was performed by two different methods: When the sample was scanned under high pressure, a linear periodic profile was produced by microploughing.³ When the sample was moved in a programmed way by using lithography software scratches were produced.

In these temperature-dependent studies the samples were annealed in a furnace the inner surface of which was lined with Al_2O_3 . SFM analyses of the surface profiles were performed under ambient conditions. The nanoscale gold powder was prepared by the inert gas condensation^{5,6} and the nanocrystalline sample, by compressing the powder at a pressure of 1 GPa in a He atmosphere.⁷

Before structuring or analyzing the samples by SFM, the surfaces were cleaned by Xe ion sputtering, and the cleanliness was analyzed by Auger electron spectroscopy (AES). After SFM analysis in air and annealing of the samples, no significant surface contamination was detected by AES. Adsorbates present during the SFM experiments in air, such as a water film, did not disturb the measurements in any noticeable degree! On the other hand, a sample with a grid produced by lithography showed a pronounced carbon contamination and a greatly reduced surface diffusion.

III. BACKGROUND

The thermal stability of nanostructures is determined mainly by surface diffusion and grain boundary diffusion and is influenced also by structural phase transitions such as roughening and surface melting.^{8,9} Also, external impacts, such as stress and electromigration must be considered. In this study, the temperature dependence of the surface structure of poly- and nanocrystalline gold was analyzed, by way of example, for local measurement of the temperature depen-

^{a)}Electronic mail: peter.blanckenhagen@imf.fzk.de

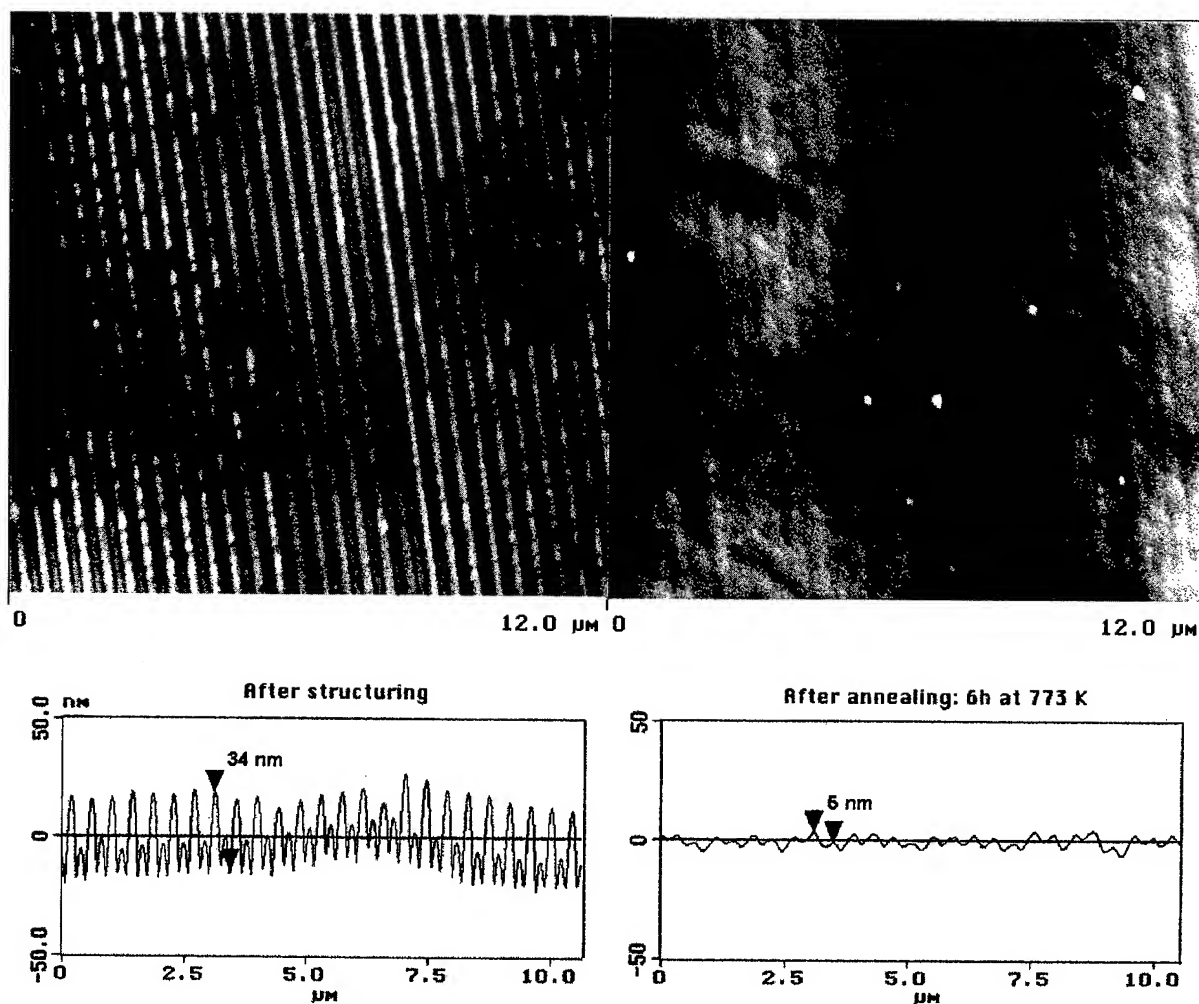


FIG. 1. SFM images and averaged profiles of a grid produced on the surface of polycrystalline gold by microplothing. (left) Immediately after production, (right) after annealing (for 6 h at 773 K).

dence of surface self-diffusion constants, of grain sizes and shapes, and of the ratio of the free energy on the surface and at the grain boundaries, respectively.

Surfaces diffusion studies analyzing micrometer-scale one-dimensional periodic profiles were performed in the past by laser beam interferometry.¹⁰ The application of SFM offers several advantages over that technique:

- (1) The analysis can be performed in sample areas in the nanometer range.
- (2) The analysis is much more sensitive and, hence, allows much lower diffusion constants to be measured, e.g., at low temperatures.

The decrease of the amplitude of a one-dimensional periodic profile as a function of temperature is furnished by the theory of Mullins *et al.*,⁴ and the references in that paper. When volume diffusion is negligible, it is given by the equations

$$a(t) = a_0 \exp(-c_4 k^4 t) \quad (1)$$

with

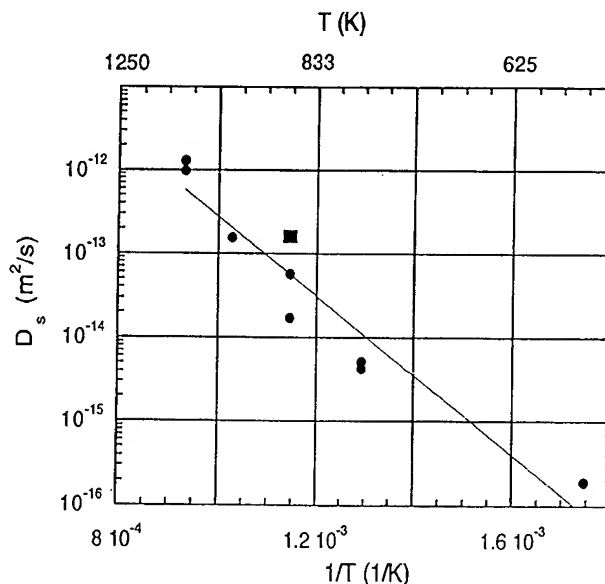


FIG. 2. Temperature dependence of the self-diffusion constant on the surface of polycrystalline gold. The square point results from a grain grooving analysis. The continuous line corresponds to a least-square fit according to Eq. (3).

$$c_4 = \frac{c_0 D_s}{k_B T} \gamma_s \Omega^2, \quad (2)$$

where k is the wave vector of the periodic profile; D_s , the surface self-diffusion constant; c_0 , the number of atoms per unit area; Ω , the atomic volume; γ_s , the surface free energy. For calculating the surface self-diffusion constants for polycrystalline Au below, the surface free energy was taken from the literature as $\gamma = 1.363 \text{ J/m}^2$.¹¹ The γ data for metal surfaces in Ref. 11 are semiempirical estimates for $T=0$ averaged over various surface orientations. No experimental temperature-dependent data for γ are available. Compared to the strong variation of D_s with the temperature, γ may be assumed to be constant.¹² It is obvious from Eq. (1) that the measurement effect can be strongly enhanced by increasing k . The activation energy, ΔE , for the diffusion processes can be derived from temperature-dependent data by applying Arrhenius' law:

$$D = D_0 e^{-\Delta E/k_B T}. \quad (3)$$

According to the theory of Watkins *et al.*,^{4,13} the surface grain boundary angle, Φ , is related directly to the ratio of the free energies of the grain boundary, γ_{GB} , and of the surface γ_s by

$$\sin \Phi = \frac{\gamma_{GB}}{2\gamma_s}. \quad (4)$$

Measuring Φ with a SFM and using a known value for γ_{GB} furnishes the γ_s needed for calculating the diffusion constant by Eq. (1) and Eq. (2), or vice versa.

IV. RESULTS AND DISCUSSION

A. Measuring the self-diffusion constant on the surface of polycrystalline gold

A detail of the periodic profile produced with the SFM and used for diffusion analysis is shown in Fig. 1. It should be noted that no grain boundary is seen in this area. After thermal treatment (annealing for 6 h at 773 K), the profile amplitude is greatly reduced due to surface diffusion. The wavelength, λ , of the original profile was 430 nm, and the maximum amplitude was 34 nm. After annealing, $\lambda \sim 404 \text{ nm}$ is measured, the difference to the original value indicating that the profiles were not ideal. However, the maximum amplitude is reduced to 5 nm by surface self-diffusion. D_s was derived from the time and temperature dependence of amplitudes applying Eqs. (1) and (2).

The temperature dependence of D_s (Fig. 2) was derived from such data producing the mean activation energy of $\Delta E_s = 0.94 \text{ eV}$. This value must be compared with the activation energy given in the literature¹⁴ for the volume diffusion of $\Delta E_v = 1.96 \text{ eV}$. The diminished interatomic interaction on the surface makes ΔE_s smaller than ΔE_v .

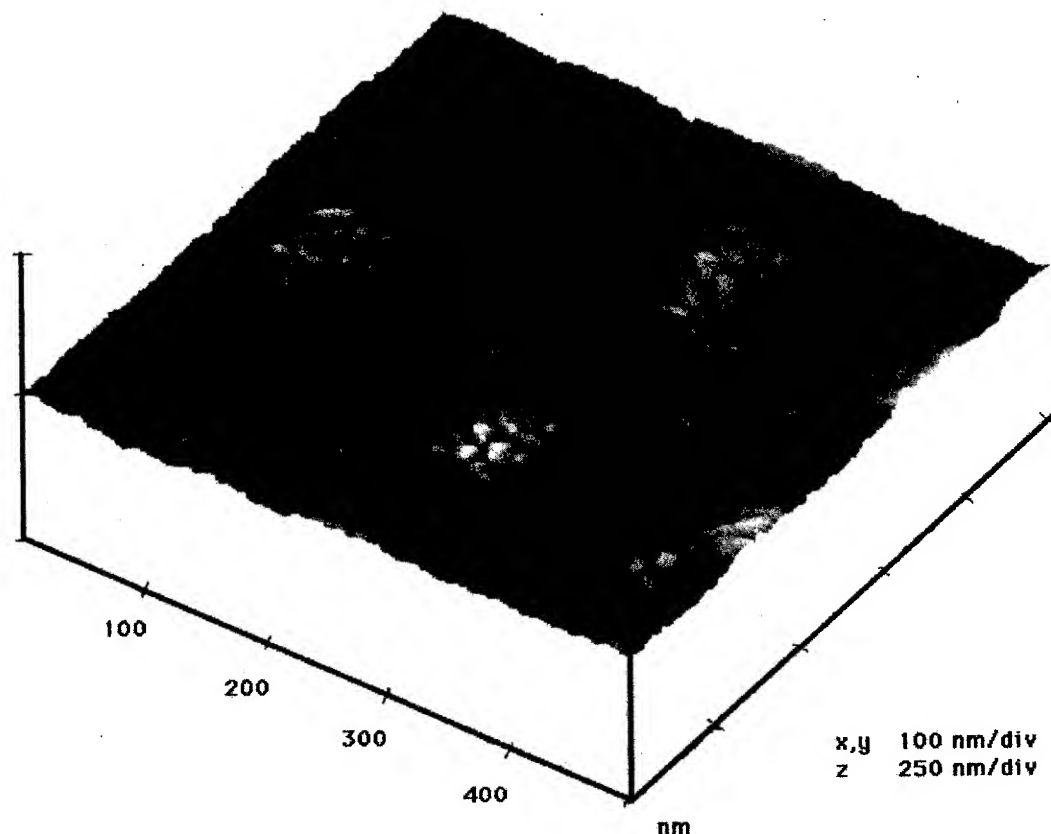


FIG. 3. Topography of the surface of nanocrystalline Au after storage for 13 months in air at room temperature.

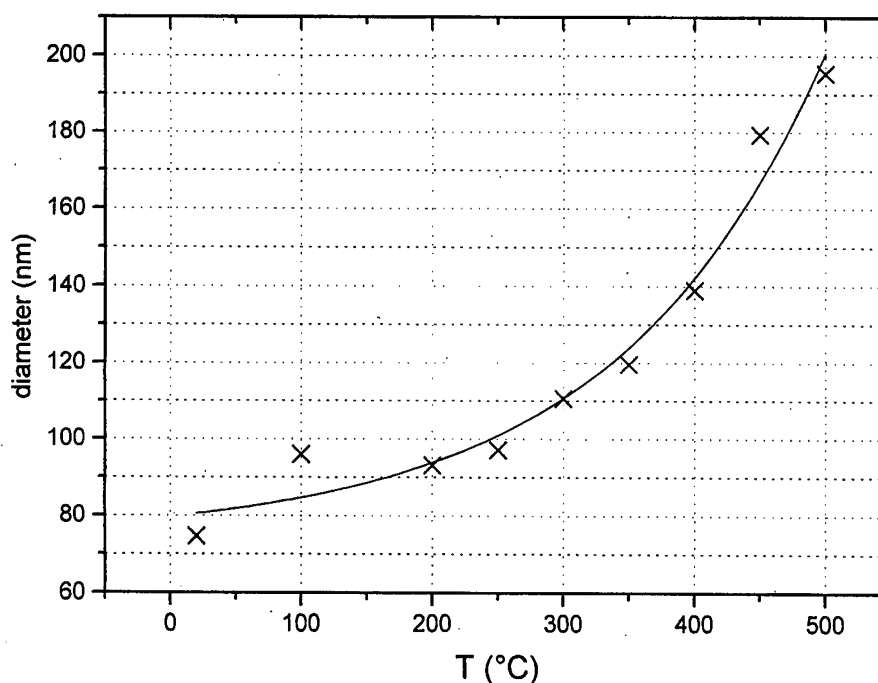


FIG. 4. Average grain size as measured with the SFM as a function of annealing temperature. Annealing time: 1 h at each temperature. The continuous line represents a fit with an assumed exponential growth law.

These results indicate the potential of diffusion studies by SFM. For other samples, and for more general application, measurements should be performed in UHV in order to prevent surface contamination, which may be non-negligible for materials other than gold, from affecting the results.

B. Temperature-dependent analysis of the surface structure of nanocrystalline gold

This explorative study proves the usefulness of SFM in temperature-dependent investigations of the surface of nanostructured material. Grain sizes, scratch broadening, surface grain boundary angles, and shapes of grains can be derived. The grain structure of a sample, which had not been treated thermally is shown in Fig. 3 by way of example. Besides small nanoscale grains, a superstructure with grains of approximately 80–200 nm diameter and with clear-cut boundaries were detected. The coarse-grained structure was studied as a function of temperature: the resultant mean grain diameter is shown in Fig. 4 as a function of temperature. Above some 400 °C a steep increase in grain size is observed, which is indicative of the transition from metastable to a stable grain structure. The observed temperature dependence of grain size rises exponentially; however, no theory describing this behavior has as yet been proposed.

When neighboring grains are in an equilibrium, the boundary is formed by two intersecting planes, and the angle Φ between the planes is given by Eq. (4). The results of a corresponding SFM analysis are shown in Fig. 5 with perpendicular sections through a grain boundary measured after annealing at 200 °C. The predicted grain boundary shape can be recognized, and the derived value of Φ corresponds to

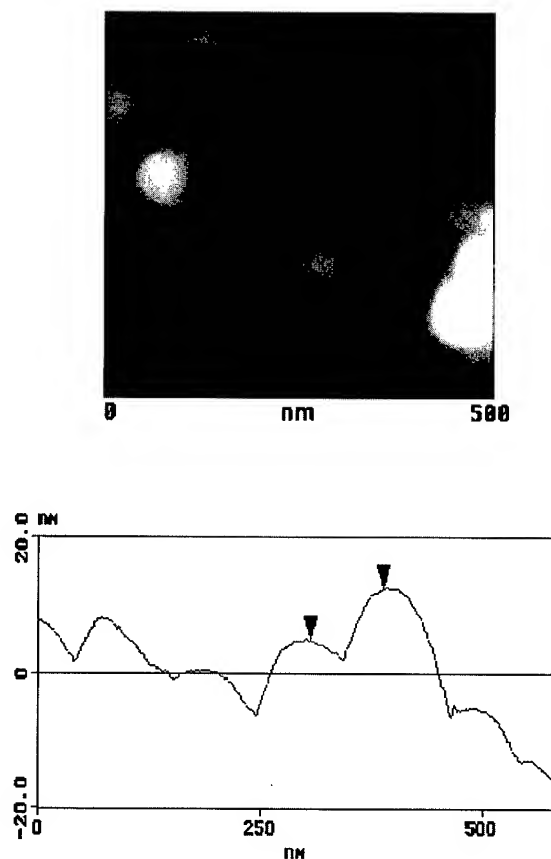


FIG. 5. SFM images of grain boundaries on the surface of nanocrystalline Au and height profile perpendicular to a boundary between two grains (after 2 h of annealing at 200 °C).

$\gamma_{GB}/\gamma_s \sim 0.3$. A similar analysis of polycrystalline gold at room temperature produced 0.1 for this ratio. For nanocrystalline Au, neither the surface free energy nor the grain boundary free energy is known. With the γ_s value for polycrystalline Au as given above γ_{GB} would be $0.2 \times \gamma_s$.

We have not yet been able to observe the grain boundary grooving effect⁴ on the nanocrystalline Au surface by SFM analysis, nor have we succeeded in structuring individual grains for measurement the surface diffusion constants.

Scratch broadening, on the other hand, was observed with rising annealing temperature, as previously also for polycrystalline gold surfaces.² In this case, however, many grain boundaries cross the scratch and there is as yet no theory by which the diffusion parameters could be extracted from the scratch broadening data.^{4,15} Consequently, small scratches on single grains need to be studied to determine the surface diffusion constant for this case.

V. CONCLUSION

Applications of the SFM for quantitative analysis of surface diffusion on Au-surfaces and related phenomena, such as grain growth on the surface of nanocrystalline Au, have been presented.

Structuring as well as subsequent temperature-dependent analysis were performed with the SFM. This technique allowed studies to be performed with high lateral resolution

and, in the case of diffusion analysis, high sensitivity not previously attained by conventional methods.

ACKNOWLEDGMENT

The authors would like to thank B. Günther for preparing the nanocrystalline gold samples.

¹R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy: Methods and Applications* (University of Cambridge, Cambridge, 1994).

²H. Göbel and P. von Blanckenhagen, *Surf. Sci.* **331–333**, 885 (1995).

³H. Göbel and P. von Blanckenhagen, *J. Vac. Sci. Technol. B* **13**, 1247 (1995).

⁴W. W. Mullins, *Metall. Mater. Trans. A* **26**, 1917 (1995).

⁵*Fundamental Properties of Nanostructured Materials*, edited by D. Fiorani and G. Sberveglieri (World Scientific, Singapore, 1994).

⁶H. Gleiter, *Nanostruct. Mater.* **6**, 3 (1995).

⁷B. Günther, Fraunhofer Institut für Angewandte Materialforschung.

⁸P. von Blanckenhagen, *Ber. Bunsenges. Phys. Chem.* **98**, 312 (1994).

⁹W. Selke and P. M. Duxbury, *Z. Phys. B* **94**, 311 (1994).

¹⁰H. P. Bonzel, *Landolt-Börnstein, Zahlenwerte und Funktionen aus Naturwissenschaften und Technik, Neue Serie, Gruppe III: Kristall und Festkörperphysik*, edited by H. Mehrer (Springer, Berlin), Vol. 26, Chap. 13, pp. 717–747.

¹¹D. Frenken, B. Pluis, and J. van der Veen, *Surf. Sci.* **239**, 282 (1990).

¹²U. Brenner and H. Bonzel, *Surf. Sci.* **273**, 219 (1992).

¹³G. L. Bailey and H. C. Watkins, *Proc. Phys. Soc. London* **63**, 350 (1950).

¹⁴I. Kaur and W. Gust, *Landolt-Börnstein, Zahlenwerte und Funktionen aus Naturwissenschaften und Technik, Neue Serie, Gruppe III: Kristall und Festkörperphysik*, edited by H. Mehrer (Springer, Berlin), Vol. 26, Chap. 12, pp. 630–646.

¹⁵M. Zinke-Allmang, *Scanning Microsc.* **4**, 523 (1990).

Scanning tunneling microscopy modification of Ag thin films on Si(100): Local rearrangement of the Si substrate by Ag/Si eutectic phase formation

U. Memmert, U. Hodel, and U. Hartmann^{a)}

Institute of Thin Film and Ion Technology, Forschungszentrum Jülich, D-52425 Jülich, Germany

(Received 10 September 1996; accepted 17 March 1997)

The melting, eutectic alloy formation, and evaporation, induced locally by a scanning tunneling microscope, was studied for 20 nm Ag films, deposited on hydrogen-terminated Si(100) surfaces. The Ag thin film can be locally rearranged or evaporated with a lateral resolution of 100–150 nm. For long interaction durations 50–70 nm wide and 30 nm deep, grooves could also be cut into the Si substrate. The modification mechanism can be explained by a model involving local melting, alloy formation, and evaporation. © 1997 American Vacuum Society. [S0734-211X(97)02204-X]

I. INTRODUCTION

The high lateral resolution of scanning tunneling microscopes (STMs) has also stimulated interest in the utilization of these instruments for controlled modifications of surfaces and thin films at an atomic scale^{1,2} as well as at a mesoscopic scale.³ In order to investigate the ability of the STM to be used as a modification tool in the processing of metallization layers in nanoelectronic devices, we have recently started a systematic study of the local melting and evaporation of Ag thin films on Si(100) for various film thicknesses, grain orientations, and grain sizes. The aim of this study is to determine if it is possible to modify not just the topmost layer of a sample, but whether or not metallic thin films can be modified over their entire film thickness as it would be needed for, e.g., the cutting of a metallic contact line in the STM-based production of nanometer-scale devices. In the first experiment performed on 20-nm-thick oriented Ag(111) films on Si(100) (2×1) substrates, it was found that it is possible to locally melt the Ag grains below the tip.⁴ The molten material is drawn towards the tip and, upon moving the tip laterally over the surface during the modification, continuous lines of molten and subsequently shock-freezed material can be written into the film. The evaporation of Ag was found to be slow such that no cutting through the film appeared to be possible. In the present article we report on modification experiments which were performed on 20-nm-thick polycrystalline Ag films, deposited on hydrogen-terminated Si(100) surfaces. For these films the sizes of the individual grains are significantly smaller. The data now show that significant local Ag evaporation can be achieved, which enables STM cutting through the entire thin film. The process can even result in a significant rearrangement of the Si substrate surfaces on a scale reaching down to ~50 nm.

II. EXPERIMENT

The experiments were performed in an ultrahigh vacuum (UHV) STM/scanning electron microscope (SEM) system

described previously.⁵ Si(100) samples were cleaned in acetone and methanol in an ultrasonic bath and were subsequently etched in 50% HF for 30 s. This etch removes the native oxide layer without attacking the Si substrate and leaves the surface in a hydrogen passivated state which remains resistant against oxidation for weeks, even under atmospheric conditions.^{6,7} The samples were transferred into UHV and used as substrates for Ag growth without any further treatment. Ag was evaporated from a Knudsen cell onto the substrates held at 300 K such that a growth rate of 0.33 nm/min was achieved and that a 1 h growth period resulted in a 20 nm film. Auger electron spectra from such films showed mainly Ag lines, but the persistence of the Si line points to quite an incomplete covering of the substrate by the film. Low energy electron diffraction displayed a largely diffuse distribution. These observations point to a growth of polycrystalline Ag grains in either a Vollmer–Weber or Stranski–Krastanov mode. STM imaging of the as-grown films displays a rather homogeneous distribution of Ag grains with a lateral extent of ~45 nm. The grains in general show a variety of different crystal facets implying that they are already the result of a coalescence of smaller differently oriented nanocrystalline grains. This observation is sustained by a STM investigation of the film morphology for various film thicknesses between 5 and 70 nm.

Before starting a modification the tip was withdrawn from the surface, which was achieved by adding a high signal to the output of the tunnel-current preamplifier. This causes the feedback loop to contract the scanner completely and remove the tip from the surface by typically 1 μm. In this situation the tip-sample bias was changed to 5 V and the current set point of the feedback electronics was adjusted to 50 nA. In addition, the feedback time constant of our electronics was decreased to a value which would result in feedback oscillations for ordinary tunneling conditions. To induce modification the added current signal was removed. This provokes the feedback electronics to dilate the scanner until a sufficient tip-sample current is detected. For point modifications the tip-surface distance was then feedback-controlled at a constant location for durations between a few seconds and several minutes. To write lines the tip was moved across the

^{a)}Also at: Institute of Experimental Physics, University of Saarbrücken, D-66041 Saarbrücken, Germany.

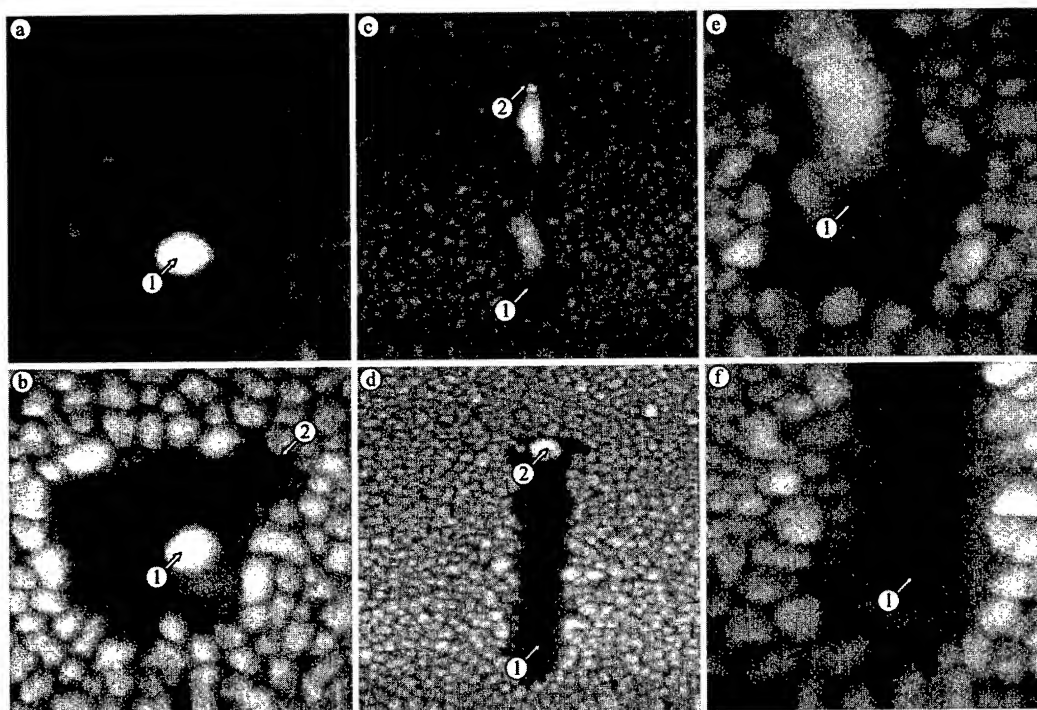


FIG. 1. STM images taken after modification of 20 nm Ag films on Si(100)-H surfaces. Modifications were performed at a junction voltage of 5 V and a current of 50 nA. (a) 500 nm \times 500 nm image showing the result of a 5 s point modification. (b) 500 nm \times 500 nm image showing the result of a 300 s point modification. (c) 1.4 μ m \times 1.4 μ m image of a 800-nm-long line, written with a speed of 750 nm/min. (d) 1.4 μ m \times 1.4 μ m image of a 800-nm-long line, written with a speed of 75 nm/min. (e) Zoom (450 nm \times 450 nm) into the lower part of the area displayed in (c). (f) Zoom (450 nm \times 450 nm) into the lower part of the area displayed in (d).

surface after an approach at 750 and 75 nm/min along a distance of 800 nm. In order to end the modification process the signal was again added to the preamplifier output and, as a consequence, the tip was retracted by the feedback electronics. Images of the modifications were subsequently taken in a constant current mode, using tunneling voltages between 0.5 and 1 V and tunneling currents between 50 and 100 pA. The data are displayed in top view representation with a linear gray scale covering a height range of 65 nm for all images presented here.

III. EXPERIMENTAL RESULTS

During the modification process, strong tip-sample oscillations with amplitudes around 100 nm were detected. This is caused by a local production of a liquid metal phase underneath the tip, which frequently leads to the creation of a liquid metal neck between the tip and the sample, also accompanied by an ohmic contact. The oscillations occur because such a ohmic contact always causes the tip to retract until the neck breaks, followed by repeated tip-sample approaches, neck formations, and retractions.⁴ The consequence of inducing a local modification over a period of 5 s at a given spot is visible in the image (500 nm \times 500 nm) displayed in Fig. 1(a). Most of the image shows the grainy structure of the polycrystalline Ag film. However, in the center where the tip has approached during modification (marked by a 1 in the image) an \sim 60-nm-high hillock with a base diameter around 100 nm is visible. This protrusion is

surrounded by a 150–200-nm-diam area, where the Ag is removed and the Si substrate is exposed. This result exhibits clear similarities to the structures which have been created by STM modification processes on oriented Ag(111) films,⁴ for modification of Au surfaces⁸ or also on surfaces of metallic glasses,⁹ and can be understood in terms of a local melting of the Ag film underneath the tip, followed by an attraction of the liquid metal to the tip. The exposed substrate surface shows a quite low roughness with only a few islands of less than 1–2 nm height which are probably remaining droplets of Ag. After such modifications we found quite often that the tip (electrochemically etched W wire) had changed its imaging capabilities, indicating that some amount of material may have been transferred between the tip and the surface. Since the modification process sometimes resulted in very sharp and sometimes very blunt tips, one has to assume that occasionally Ag from the film has been transferred to the tip or that Ag, being present at the apex of the tip from a previous experiment, has been transferred to the film. Upon comparing the results of several experiments performed under identical conditions, we never found indications that this may have caused the transport of volumes of material being significant in comparison with those volumes of modified material in the film. In the present case a rough comparison of the volume of the protrusion with the substituted film volume in the modified area indicates that only insignificant amounts of Ag have been lost or gained during the modification process.

In order to investigate the effect of longer durations on the modification process and to check the possibility of inducing significant Ag evaporation, we performed a similar experiment with the substantially extended modification duration of 300 s. The resulting topography ($500 \text{ nm} \times 500 \text{ nm}$) is displayed in Fig. 1(b). Similar to the image shown in Fig. 1(a), the outer parts of the imaged area display the grainy structure of the unmodified film and the location where the tip had approached the surface (indicated by 1 in the image) exhibits a protrusion. However, apart from these similarities, the resulting topography is quite different from the topography found before. The central protrusion is not as high as the former one, reaching only up to $\sim 40 \text{ nm}$ above the substrate level. The total area within which the Ag film was affected has increased by a factor of 2.1 and at the upper left of the protrusion the substrate has been significantly affected as well. The resulting hole extends 40 nm down into the Si substrate. Most of the exposed substrate area displays a roughness of up to 5 nm in amplitude, being the result of a high density of $\sim 10\text{--}20 \text{ nm}$ wide and 5 nm high islands. Only a rather remote area, marked by 2 in the image, shows a lower density of islands, similar to what was also visible in Fig. 1(a). In addition, for the extended modification time, the amount of material present in both the central protrusion and the small islands on the substrate is much lower than the amount of material which must have been present in the substrate and the film before modification. One finds that the remaining material is just sufficient to account for the volume of material being removed from the hole in the Si substrate, while the entire volume of material, which formerly constituted the Ag film, appears to be missing. The clear implication is that during elongated modifications not only the film and the substrate is affected, but also a significant loss of material takes place.

The process of tip-induced local melting and evaporation can also be used to write lines or other, more complicated, patterns into the film. The image ($1.4 \mu\text{m} \times 1.4 \mu\text{m}$), displayed in Fig. 1(c), was taken at a position where a line was written, employing a writing speed of 750 nm/min . The tip was approached in the lower part of the imaged area at the position marked by 1 and was then moved upwards along a distance of 800 nm until being withdrawn at the position marked by 2. Surrounded by the grain structure of the unmodified Ag film, a solid line is visible, connecting the points of approach and withdrawal. Similar to what already was found for the oriented Ag grains on Si(100) (2×1),⁴ this continuous line results from locally melting Ag underneath the tip. The liquid is attracted by the tip and left on the surface during the tip motion. The height of the central line measured from the substrate level exceeds at several positions the original film thickness of 20 nm . It appears like a viscous liquid spread out over the substrate. Its surface is quite smooth and has a rounded appearance. Crystal facets, as resolved on the Ag grains of the film and which would be a hint for the presence of a recrystallization process during cool down, are not visible on its surface. From this it is concluded that the very fast cooling process after/during

writing has left the material in an amorphous state. Similar to the point modification experiment, the central line is surrounded by a depletion zone. A rough comparison of the total amount of material in the central line with the total amount of material, which was present in the entire modified area before modification, clarifying that evaporation of Ag did again not play a significant role.

Similar to the effect of extended modification periods during the point modification experiments, significant evaporation is also found in the line writing experiments when moving the tip across the surface at a speed of 75 nm/min (i.e., reduction by a factor of 10). The result of such an experiment is displayed in Fig. 1(d) ($1.4 \mu\text{m} \times 1.4 \mu\text{m}$). Again, the surface was approached at the position marked by 1 and the tip was withdrawn at the position marked by 2. One now finds that the central feature is a groove, reaching down to 30 nm into the substrate. Only at the location of withdrawal did we find a protrusion, extending only $\sim 30 \text{ nm}$ above the substrate level. The width of the modified area varies between 150 and 250 nm and the width of the central notch is as narrow at $50\text{--}70 \text{ nm}$.

The characteristic features of the resulting topography become clearer in images ($450 \text{ nm} \times 450 \text{ nm}$) taken after zooming into the lower part of the modified areas. Figure 1(e) displays an area from the lower part of Fig. 1(c) (750 nm/min) and Fig. 1(f) shows the lower part of the area displayed in Fig. 1(d) (75 nm/min). In both cases, the point of initial tip approach is again marked by a 1. The most interesting comparison between these two images concerns the roughness of the areas from which the Ag was removed. While the substrate, exposed during the fast writing process, shows a low roughness with only a few $1\text{--}2\text{-nm}$ -high islands, in the case of the slow modification the entire modified area is covered by a dense distribution of up to 5-nm -high and $10\text{--}20\text{-nm}$ -wide islands, which at closer inspection also extend down to the bottom of the groove. For this and similar images we found that only at the outermost parts of the modified area does the density of these islands appear to decrease slightly, similar to what was already also visible for the long-duration point modification experiment [Fig. 1(b)].

IV. DISCUSSION

The experimental data clearly show that it is possible to use the tip of a STM to locally melt and even evaporate 20 nm Ag films on a Si substrate. For the local melting process we found that we were able to achieve a width of the solid line of $100\text{--}150 \text{ nm}$. A more interesting observation is that for extended writing periods or low writing speed we were not only able to locally melt and evaporate the Ag film, but that even a significant cutting into the Si substrate was achievable. This observation seems quite puzzling, since the bulk melting temperatures of Ag and Si differ significantly (1234 K for Ag and 1685 K for Si) and also the vapor pressures of both materials are quite different in the entire regime of interesting temperatures with the Ag vapor pressure exceeding the Si vapor pressure by typically four to five orders of magnitude.

The fact that the Ag film was locally molten below the tip could point to local temperature above 1200 K being reached at least for some portions of the total modification period. However, due to an interdiffusion of Si into Ag the temperature of liquid phase formation may even be lowered, since this would result in the creation of an eutectic alloy with a liquidus of 1125 K.¹⁰ In addition, the melting temperature of a thin film or of such small particles as the Ag grains in the film can be substantially lowered as compared to a bulk material.^{11,12} From our data, the achieved temperatures appear to be just sufficient to result in Ag evaporation for the slower speed, respectively for the longer modification period [compare Figs. 1(b), 1(d), and 1(f)]. Already for the faster speed or the shorter modification duration no significant Ag evaporation was observed [compare Figs. 1(a), 1(c), and 1(e)]. Taking the significant difference in vapor pressure and melting point for Ag and Si into account, one finds that local temperatures, being just sufficient to evaporate Ag, would not be enough to melt or even evaporate clean Si. A control experiment, performed on a hydrogen-terminated Si surface without the presence of a Ag film, also clarified that no similar modifications could be introduced to the bare substrate. This is also in good agreement with previously published experiments on the tip-induced modifications of hydrogen-terminated Si surfaces, which resulted only in hydrogen desorption without mesoscopic-scale rearrangements of the Si, although the junction voltages were quite similar to the parameters used here.^{13,14}

The experiments, however, show that for a longer duration or a lower writing speed the substrate is significantly affected in the present case of Ag thin films on Si. It is even possible to cut deep grooves in the Si substrate. It should therefore be assumed that the modification of the Si substrate is mediated by the presence of Ag on the surface via the formation of an eutectic alloy. However, although this can account for a liquefaction of Si at rather low temperatures, it cannot result in Si evaporation with sufficiently high rates. The activity coefficients in a liquid Ag/Si mixture do not differ more than 20% from unity,¹⁰ i.e., the vapor pressure of Si above an Ag/Si solution does not significantly deviate from Raoult's law. Therefore the Si vapor pressure will not be increased due to the eutectic phase production. Instead, it will even be lower above the liquid mixture than it would be above clean Si, since the Si content in the solution is less than 100%.

From these considerations it becomes clear that a qualitative model explaining our experimental observations has to take into account significant Ag evaporation for longer modification durations on one hand and should, on the other hand, explain the significant rearrangement of the substrate without assuming evaporation of Si. Figure 2 displays schematic sketches illustrating the processes we propose will take place during the modification. A temperature around 1100–1200 K is reached during modification and the Ag film is molten. In the center, where the temperatures are the highest, the Si substrate is also affected and an Ag/Si solution is formed leading to an intermixing of both materials within the entire

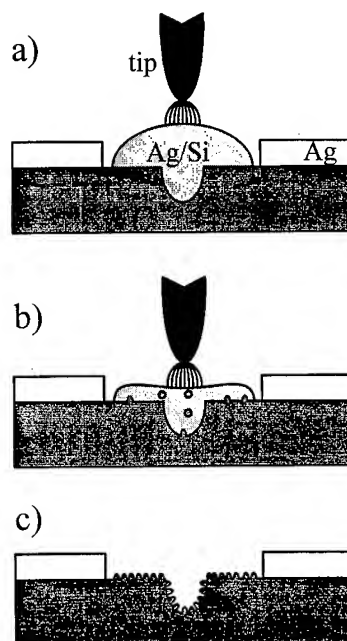


FIG. 2. Schematic illustration of the modification process: (a) Local heating and intermixing of Ag and Si by the STM tip. (b) Evaporation of Ag and redeposition of Si. (c) Final structure where the Ag was removed and the Si was rearranged.

liquid volume. This situation, which is depicted schematically in Fig. 2(a), is frozen in when the tip is withdrawn again after quite short times or when it is moved over the surface at high speed. The molten and subsequently frozen material then gives the visual impression of a viscous liquid in the images (compare Fig. 1). Upon extending the interaction duration at each location, Ag desorption becomes effective. The evaporation of Ag will lead with time to an increasing enrichment of Si in the liquid solution, which then gradually results in the redeposition of Si at all those locations on the substrate which are covered by the liquid solution. This situation is depicted in Fig. 2(b). Finally, after evaporation of all Ag, the modified area will remain in a state as it is found in the images displayed in Figs. 1(d) and 1(f) and as it is sketched schematically in Fig. 2(c). Now the Ag is removed completely from the modified area and some Si is moved from the center, where the temperatures had previously reached the highest values, to the entire area on which the interface of the liquid solution and the substrate existed previously. This creates a notch in the center, although no substrate material has been evaporated. The grains created during this Si redeposition are clearly visible as the spatially restricted grainy structure in Figs. 1(b) and 1(f) and extend in height up to 5 nm above the former substrate level. Only in the outer parts of the modified area, from which Ag has been drawn into the center and which were not wet by the liquid, does the bare substrate with a much lower density of islands remain exposed. Such a situation is visible at the location, marked by 2 in Fig. 1(b).

V. CONCLUSION

Operating a STM at a rather high junction voltage of 5 V permits the modification of Ag thin films on Si(100) surfaces by local heating underneath the tip. The heating is sufficient to melt the Ag film and to locally create a liquid Ag/Si solution. For short interaction times, the liquid solution remains on the surface and continuous lines with a width around 100 nm can be written into the grainy film. For relatively long interaction times, the Ag can be removed completely from the surface and even the Si substrate can be significantly rearranged due to the creation of an eutectic phase and the formation of a liquid Ag/Si solution, followed by a redeposition of Si during evaporation of the Ag. Using this method, 50–70-nm-wide and 30-nm-deep grooves can be cut into the Si substrate.

¹D. M. Eigler and E. K. Schweitzer, *Nature* **344**, 524 (1990).

²T. A. Jung, R. R. Schlittler, J. K. Gimzewski, H. Tang, and C. Joachim, *Science* **271**, 181 (1996).

³R. Wiesendanger, *Jpn. J. Appl. Phys.* **1** **34**, 3388 (1995).

⁴U. Hodel, U. Memmert, and U. Hartmann, *Phys. Rev. B* **54**, 17888 (1996).

⁵U. Memmert, U. Hodel, and U. Hartmann, *Rev. Sci. Instrum.* **67**, 2269 (1996).

⁶U. Neuwald, A. Feltz, U. Memmert, and R. J. Behm, *Appl. Phys. Lett.* **60**, 1309 (1992).

⁷M. Morita, T. Ohmi, E. Hasegawa, M. Kawakami, and M. Ohwada, *J. Appl. Phys.* **68**, 1272 (1990).

⁸R. Taylor, R. S. Williams, V. L. Chi, G. Bishop, J. Fletcher, W. Robinett, and S. Washburn, *Surf. Sci.* **306**, L534 (1994).

⁹U. Stauffer, L. Scandella, and R. Wiesendanger, *Z. Phys. B* **77**, 281 (1989).

¹⁰Landolt-Börnstein, *New Series*, edited by O. Madelung (Springer, Berlin, 1991), Vol. IV/5a, pp. 100–101.

¹¹N. T. Gladkich, R. Niedermayer, and K. Spiegel, *Phys. Status Solidi* **15**, 181 (1966).

¹²É. L. Nagaev, *Sov. Phys. Usp.* **35**, 747 (1992).

¹³R. S. Becker, G. S. Higashi, Y. J. Chabal, and A. J. Becker, *Phys. Rev. Lett.* **65**, 1917 (1990).

¹⁴J. W. Lyding, G. C. Abeln, T. C. Shen, C. Wang, and J. R. Tucker, *J. Vac. Sci. Technol. B* **12**, 3735 (1994).

Novel lithography and signal processing with water vapor ions

H. W. P. Koops

Deutsche Telekom AG, Technologiezentrum, D-64295 Darmstadt, Germany

E. Dobisz

Naval Research Laboratory, Washington, DC 20375

J. Urban^{a)}

Institut für Angewandte Physik der Technischen Hochschule Darmstadt, D-64289 Darmstadt, Germany

(Received 12 September 1996; accepted 31 March 1997)

A novel model is presented for the operation of field emission sources in air. Under this condition the cathode-anode distance is less than the mean free path for the particles emitted in air. Such conditions are encountered in scanning tunneling microscope lithography in air and in novel integrated microtriode tube devices designed to operate in air. In air, the presence of a high electric field gradient, produced at the tip, will cause the polar water molecules coating the surface to align with the field and move to the high field region. As a result, a Taylor cone of water, like that as it is observed with liquid metal ion sources, will form. The high electric field gradient at the Taylor cone, enhanced by the high dielectric constant, will produce a very finely focused beam of ions. We examine the case of water ion formation. With a positive tip bias, hydronium ions will form at electric fields of 1 V/nm. This is a factor of 4 smaller than that required to produce the field emission with electrons. The electric field distribution and the hydronium ion trajectories from a metal tip with a high permittivity material cone are calculated. The spot size of the ion beam on the substrate was 2.9 nm. A model for lithography, based on water ions, is presented. The model is applied to the design of a microtriode device for signal processing. The three-dimensional fabrication of such a device by e-beam induced chemical vapor deposition is demonstrated. © 1997 American Vacuum Society. [S0734-211X(97)03804-3]

I. INTRODUCTION

Scanning tunneling microscope (STM) lithography is performed with a tip-sample bias $> 3-4$ V. In this mode the tip operates in the field emission mode, rather than in the tunneling mode. There have been several reports on STM lithography performed in air. It was found that the humidity and the tip material strongly affect the lithographic performance of the proximal probe.¹ Both oxidation of the substrate and oxidation of the tip have been observed.¹⁻³ In air the tungsten tip-sample current decreases to zero if a positive tip-sample bias is applied. Tungsten tips also function well with a negative tip-sample bias however.^{1,2} In air both imaging and lithography can be performed with a positive tip bias if noble metal tips are employed. Oxidation of hydrogen passivated silicon was observed in STM lithography using a positively biased Pt/Ir tip.⁴ Diffusion limited anodic oxidation of a silicon substrate induced by an atomic force microscope (AFM) with a negatively biased conductive tip has also been reported.³ The field induced oxide growth was found to terminate with a thickness of 8 nm at 8 V tip-sample bias. The authors of Ref. 3 propose that the oxidation process terminated when the thickness became too large to support a 1 V/nm electric field across the oxide. Such a field is required to generate OH^- ions and drive the ion migration necessary for the oxide growth.

In this article an alternative model for the oxidation process is presented. We know that in air all hydrophylic sur-

faces are coated with several molecular layers of water. The effect of the high electric field gradient on the water layer and its implications are considered. We propose that a Taylor cone of water is formed at the tip by the balanced forces that consists of the electric field acting on the permanent dipole moments of the molecules and the surface tension of the water.^{5,6} In a balanced state of internal pressure, surface tension, and tip electric field, a cone of water is formed. The highest electric field inhomogeneity is at the tip of the cone. The high field at the tip ionizes the water molecules and a beam of ionized material is emitted. An equilibrium is established between material forced to the tip by the field and ionized material emitted from the tip. Taylor-cone formation has been extensively observed in liquid metal ion sources. It produces a very fine beam of ion current, even if the tip is blunt.⁶

From field ionization mass spectroscopy, water has been reported to ionize to hydronium ions H_3O^+ at a field of 1 V/nm.⁷ In the case of a negative tip bias a similar cone must also be generated, which then would emit OH^- or H_3O_2^- ions. In either case a current of oxygen containing ions is directed to the sample. Employing hydroxyl and hydronium ions the chemical reactions with hydrogen passivated silicon and with metals can be formulated in a convincing way. We propose further that the deficient ion species shifts the pH value of the water at the tip, which can lead to chemical reactions with the tip material. A positive tip is oxidized by the hydroxyl ions, whereas a negative tip is passivated by the surplus of protons.

^{a)}Present address: Carsonweg 80, D-64289 Darmstadt, Germany.

Because water is a high dielectric material, it expels the electric field. Hence an even higher electric field gradient will form on a water Taylor cone than on a comparable radius conducting tip. The emitted ions are accelerated by a few eV energy by the tip-sample bias. In STM applications emission current ranges from pA to several nA. This process requires less power than electron field emission, where 3–4 V/nm are required.

In the arguments presented below, we propose that lithography and the operation of microtriode tubes in air are performed by water based ions. The model is used to provide alternative explanations to reported experimental results. Future experiments to test the model are outlined.

II. SIMULATION OF THE CHARGED PARTICLE OPTICS IN A STM WITH A WATER TAYLOR CONE

The potential distribution in a STM tip with a water Taylor cone was determined by a numerical simulation. The simulation was performed for a 13-nm-diam tip coated with a 1 nm layer of water, with a water cone with permittivity of 81 on top. For our exploratory work, a full cone angle of 105° was used, the same as reported for liquid metal sources. Due to its complexity, a computation of the balance forces to form a water Taylor cone is left to future work. However, for the argument and principles discussed here, the exact Taylor-cone angle is not critical. A tip with a 20 V bias was placed 20 nm from a conducting wafer that is coated with 5 nm of resist having a permittivity of 3. The resist is also coated with a 1 nm water layer. Numerical solution of the problem was obtained using Munro's electron-optic three-dimensional (EO-3D) software with an computational accuracy of 10^{-9} for the error. The simulation required 9 h on a 50 MHz 486 PC to iterate the solution on 243 000 nodes. The calculated potential distribution is shown in Fig. 1. Figure 1(a) shows the potential distribution over the entire simulated geometry. Figure 1(b) gives a magnified view of the potential distribution about the tip. The field is small in the resist and 1 V/nm at the resist surface. In Fig. 1(b) a large gradient of the axial electric field is clearly visible inside and outside the dielectric of the Taylor cone at the tip. Here the axial electric field rises to 4 V/nm. This forces the water molecules to the tip, where they are field ionized and emitted as ions.

In this simulation, the paths of emitted hydronium ions that may be used in the lithography mode were computed. The Taylor cone enhances the focusing of the emitted ions and confines the area of the emission current, as shown in Fig. 2. Traces of hydronium-ion rays as emitted from a Taylor cone with a tangential energy width of 27 meV corresponding to the thermal energy of water at room temperature are presented. The modeled geometry gives an ion beam of 2.9 nm spot size at the sample.

III. STM LITHOGRAPHY WITH WATER LAYERS

A STM with a fine tip, positively biased at 100 mV, operating in air, can produce a tip-sample current of several nA.⁴ This mode of operation allows STM imaging of con-

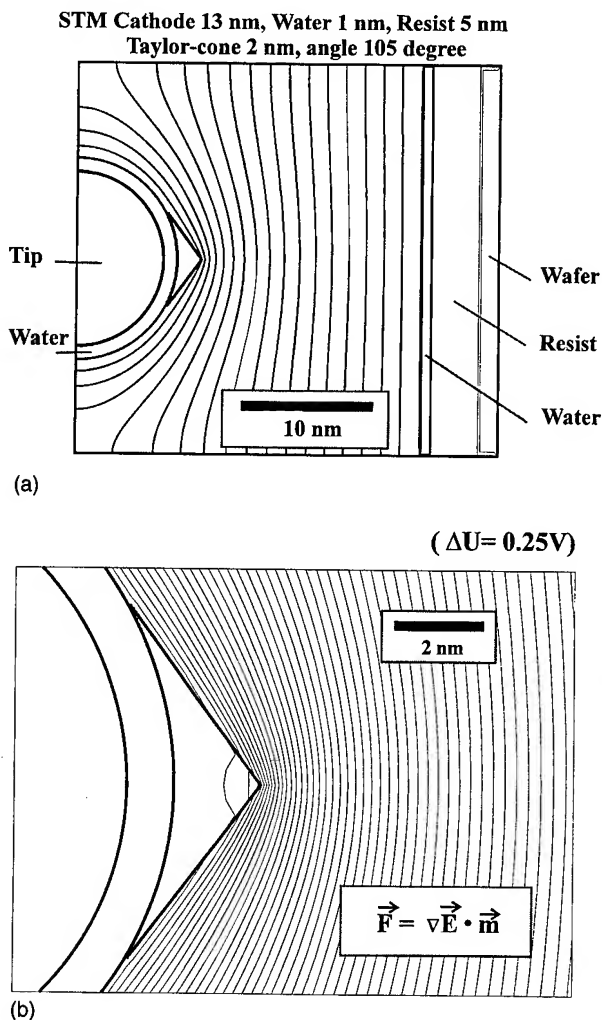


FIG. 1. (a) Potential distribution in STM with 20 nm tip-wafer distance and 20 V tip voltage. A Taylor cone of water with a permittivity of 81 is on the tip. (b) Potential distribution in higher magnification and resolution around the tip. Clearly a gradient of the axial field is visible inside the dielectric cone. This forces the water molecules to the tip, where they are field ionized and emitted as hydroxyl or hydronium ions.

ducting surfaces. Normal, nondestructive imaging of hydrogen passivated silicon is possible up to tip-sample voltages of 2.5 V. Increasing the tip-sample voltage causes surface modification. The structure generated on the hydrogen passivated Si is an oxide that can be used to mask the Si during etching.⁴

Experiments with a STM in air using cut wire Pt/Ir tips produced permanent traces that were less than 1.5 nm high but the trace width depended on the tip geometry. This can be explained by the fact that the tip geometry is imaged in the field distribution and forms a wedge-shaped Taylor cone in the case of a ridge-shaped tip. Hence, the emission site for the generation of hydronium ions is distorted. In the case of a ridge an astigmatic source is revealed from the deposited structures and could be correlated to the actual tip orientation; see Fig. 3. The tip voltage was +4.0 V. The recording speed was 200 nm/s. The image displays a height variation of 2.9 nm from black to white. The oxide thickness is 1 nm.

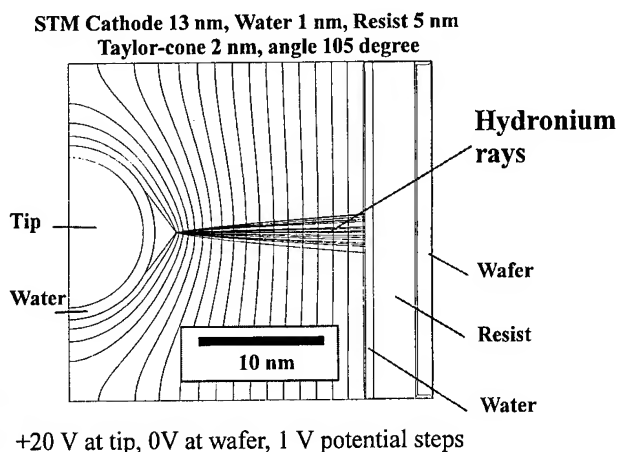


FIG. 2. Traces of hydronium ion rays as emitted from a Taylor cone with a tangential energy width of 27 meV corresponding to the thermal energy of water at room temperature.

Similar results were also reported by other authors.³

With our model, an alternative explanation of the limitation of the oxide thickness in AFM/STM lithography reported by Gordon *et al.*³ can be given. Due to the permittivity of the oxide there is no large potential drop across an oxide, however, the oxide prevents the current flux of ions emitted from the STM tip. Hydronium ions of 8 eV do not have the kinetic energy to penetrate the 8 nm silicon oxide to grow the oxide further. Therefore a positive space charge is collected at the surface of the oxide facing the tip. This reduces the hydronium emission and the oxidation process stops.

Our model can also explain the observation that a positively biased tungsten tip in air ceases to function with time. This is now understood, since the emission of the protons together with the water molecules, forming hydronium

H_3O^+ ions, will leave an increasingly basic medium at the tip. The hydroxyl ions are attracted to the tip and oxidize the tungsten tip to an extent that prohibits any current flux to the Taylor cone. The emission current goes to zero because the tip has grown a nonconductive oxide layer. This oxidation cannot occur with Pt/Ir tips. Therefore those tips work continuously in air when positively biased sufficiently to perform lithography in this mode.⁴

Our model has been employed to provide explanations for reported experimental results. Campbell¹ and Snow are in the process of constructing a N_2 atmosphere lithographic chamber in which the amount of water vapor is controlled. Lithographic results using W, doped Si, and Pt/Ir tips will provide key tests of the model.

IV. CONSTRUCTION OF A MINIATURIZED TRIODE USING ADDITIVE LITHOGRAPHY WITH ELECTRON-BEAM INDUCED DEPOSITION

The impact of the proposed model of ion emission from field emitter tips may be stronger in the field of high frequency electronics than in nanostructure fabrication. Many applications in microelectronics and telecommunications require amplifying switches operating in the upper GHz regime. Power requirements are of high interest in this application, especially if millions of sources are involved to deliver the required high currents. Using the proposed model, a miniaturized triode device was designed and fabricated with a cathode to anode distance of the order of the mean free path in air at 1013 hPa. The triode device designed has a very small cathode to extractor grid capacitance. Such devices require little or no vacuum. This is in sharp contrast to field emitter devices in vacuum microelectronics. By employing the hydronium source the power required to operate the tube could be a factor of 4 lower than in miniaturized electron triode devices because of the lower turn-on voltage.

Such microtriodes can now be constructed using additive three-dimensional electron-beam lithography.⁸⁻¹⁰ E-beam induced chemical vapor deposition can be performed in a scanning electron microscope, in which the e-beam, controlled by computer, writes a pattern. Field emitters produced by an e-beam deposition process have additional field enhancement and beam emission confinement due to the nanocrystallinity of the deposited material.¹¹⁻¹³ For triodes with electron field emitter sources operating in 1013 hPa, it is estimated that amplifying switches can operate close to 200 GHz.^{14,15} This is also true for tubes operating in humid air with the hydronium emission as a current source. The design layout and the calculated equipotential contours for a microtriode with sub-micron cathode to anode distance is shown in Fig. 4. In this design the electron or ion beam is directed parallel to the substrate surface. The new design employs a long needle as the field emitter cathode and anode, respectively. The extractor is a free-standing structure with one closed loop as a grid, which is a wirelike structure deposited by the e-beam under computer control. A micrograph of the fabricated microtriode is shown in Figs. 5(a) and 5(b) showing, respectively,

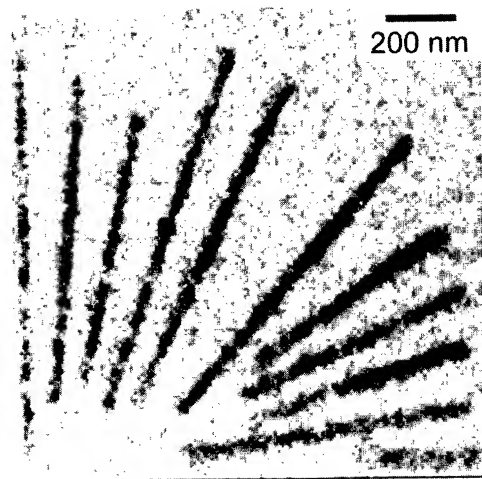


FIG. 3. Lithography with a STM in air performed with a cut wire Pt/Ir tip on hydrogen passivated silicon. The tip voltage was +4.0 V. The recording speed was 200 nm/s. The orientation dependent linewidth reveals the astigmatic current source defined by the ridge. The image displays a height variation of 2.9 nm from black to white. The oxide thickness is 1 nm.

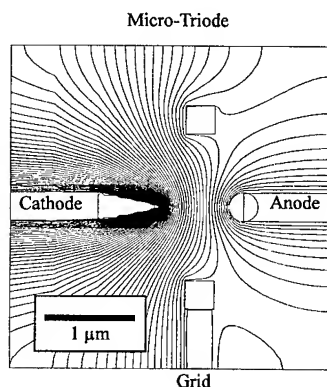


FIG. 4. Layout and calculated equipotentials in a microtriode 250 nm long.

the side view and the top view. The entire device was deposited within 6 min from adsorbed molecules of dimethyl-gold-trifluoroacetylacetonate onto a gold wiring pattern that was prefabricated with lithography and a lift-off process. To prevent corrosion the electrodes must be built from noble metals like platinum or gold. This has already been accomplished with electron-beam induced deposition. Experimental verification of the proposed operation of the triode in air is planned in the near future.

Future work will involve tests of the performance of the device in air and in a vacuum under both positive and negative bias. The results will provide critical tests of the model.

For the electron emission microtriode device a transconductance of $10 \mu\text{S}$ is expected. In the ionized water case, the source of the particles is indistinguishable. The Taylor-cone action will pump the water molecules to the tip of the cone. Hydronium ions, once emitted will neutralize at the anode and vaporize again to be attracted to the tip by the electro-

static forces. This pump will supply the amount of neutral particles required to operate the triode continuously.

V. CONCLUSIONS

We have proposed a model for field emission tip operation in air. In our model, the water layer on the tip is driven to form a Taylor cone by the interaction of the high electric field gradient and the dipole moment of the water molecules. With a positive tip bias, water will ionize to hydronium ions at a factor of 4 lower electric field than required for electron field emission. The equipotential contours and particle optics were calculated for a water Taylor cone and hydronium ion emission. It has been argued that STM lithography in air is really ion beam lithography with the oxidizing action of the water ions. These can oxidize the tip as well as the sample, depending on tip material and bias polarity. With such an ion source, it is estimated that microtriodes with $0.5 \mu\text{m}$ cathode to anode distance can possibly be operated in air. It was shown that such three-dimensional triode devices can be fabricated with e-beam induced deposition. This leads to the high expectation that water based ion emission will allow very high frequency signal processing in "Hydronics."

ACKNOWLEDGMENTS

The authors are very grateful for the excellent earlier collaboration with M. Weber, J. Kretz, and M. Rudolph in the field of electron-beam induced deposition and in the field of STM lithography with S. Bräuer and H. Pagnia of the Institut für Angewandte Physik der Technischen Hochschule Darmstadt, and with A. W. McKinnon, J. Horn, K. Hjort, and H. L. Hartnagel of the Institut für Hochfrequenztechnik der Technischen Hochschule Darmstadt. They express their appreciation to Deutsche Telekom, Technologiezentrum, for supplying the STM head in a research grant to the Institut für Angewandte Physik der Technischen Hochschule Darmstadt, where the challenge for this work was generated.

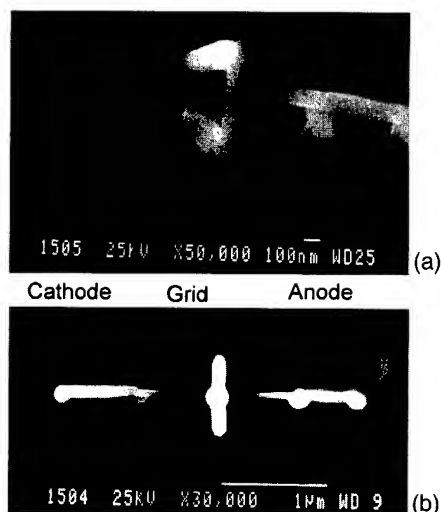


FIG. 5. (a) Side view of a fabricated microtriode with a cathode to anode distance of 700 nm. (b) Top view of the fabricated microtriode.

- ¹P. Campbell, Naval Research Laboratories, Washington, DC (private communication).
- ²E. S. Snow and P. M. Campbell, *Science* **270**, 1639 (1995).
- ³A. E. Gordon, R. T. Fayfield, D. D. Litfin, and T. K. Higman, *J. Vac. Sci. Technol. B* **13**, 2805 (1995).
- ⁴J. Urban, S. Bräuer, A. W. McKinnon, J. Horn, K. Hjort, H. Pagnia, H. W. P. Koops, and H. L. Hartnagel, *Microelectron. Eng.* **27**, 113 (1994).
- ⁵G. I. Taylor, *Proc. R. Soc. London, Ser. A* **280**, 383 (1964).
- ⁶R. G. Forbes and Z. Djuric, *Technical Digest 9 of the International Vacuum Microelectronics Conference*, 1996, p. 468.
- ⁷H. D. Beckey, *Z. Naturforsch. Teil A* **14**, 712 (1959); R. Gomer, *Field Emission and Field Ionization* (AIP, New York), p. 181.
- ⁸H. W. P. Koops, R. Weiel, D. P. Kern, and T. H. Baum, *J. Vac. Sci. Technol. B* **6**, 477 (1988).
- ⁹J. Kretz, M. Rudolph, M. Weber, and H. W. P. Koops, *Microelectron. Eng.* **23**, 477 (1994).
- ¹⁰H. W. P. Koops, J. Kretz, M. Rudolph, M. Weber, G. Dahm, and K. L. Lee, *Jpn. J. Appl. Phys.* **33**, 7099 (1994).
- ¹¹H. W. P. Koops, M. Weber, J. Urban, and C. Schössler, *Proc. SPIE* **2522**, 189 (1995).
- ¹²C. Schössler, A. Kaya, J. Kretz, M. Weber, and H. W. P. Koops, *Microelectron. Eng.* **30**, 471 (1996).
- ¹³H. W. P. Koops, C. Schössler, A. Kaya, and M. Weber, *J. Vac. Sci. Technol. B* **14**, 4105 (1996).
- ¹⁴H. W. P. Koops and C. Schössler, in *Ref. 6*, p. 458.
- ¹⁵H. W. P. Koops, German Patent Nos. 4,416,597.8 and 19,609,234.5.

Fabrication of surface nanostructures by scanning tunneling microscope induced decomposition of SiH_4 and SiH_2Cl_2

H. Rauscher,^{a)} F. Behrendt, and R. J. Behm

Abteilung Oberflächenchemie und Katalyse, Universität Ulm, D-89069 Ulm, Germany

(Received 12 September 1996; accepted 30 January 1997)

Nanostructures with lateral dimensions below 10 nm are fabricated on $\text{Si}(111)7\times7$ surfaces by scanning tunneling microscope induced dissociation of the precursor gases SiH_4 and SiH_2Cl_2 . The molecules are activated between the tunnel tip and the surface, where highly reactive species are produced by the interaction of molecules with free electrons in the tunnel gap. The molecular fragments accumulate directly below the tip and build up the nanostructure. Information on the deposition mechanism that is effective at the onset of deposition is obtained by careful evaluation of the deposition voltage threshold regime. © 1997 American Vacuum Society.

[S0734-211X(97)10704-1]

I. INTRODUCTION

Soon after its invention, the scanning tunneling microscope (STM) was utilized not only for surface imaging but also for surface modification. A number of different methods for surface modification with the STM has subsequently evolved and they are reviewed in Refs. 1 and 2. These studies include selective chemical vapor deposition from metal³⁻⁹ or semiconductor-containing chemical vapor deposition (CVD) precursors,¹⁰⁻¹² activation of hydrogen-passivated Si surfaces,^{13,14} exposure of electron-sensitive resist materials,^{15,16} manipulation of single surface atoms,¹⁷ and selective breaking of substrate¹⁸ or adsorbate bonds.¹⁹ In this study we employ the method of (local) STM-CVD which was first demonstrated in Ref. 3 and subsequently also by a number of other workers⁴⁻⁹ for the deposition of metallic nanostructures. In these experiments low energy electrons emitted from a STM tip are used to locally decompose a precursor gas. STM-CVD fabrication of structures from the semiconductor precursor SiH_4 on Si surfaces was demonstrated in Ref. 10 and subsequently also by Wong *et al.*¹¹ and deposition of Ge-containing spots on $\text{Si}(111)$ was reported from STM induced decomposition of GeH_4 .¹²

In these earlier STM-CVD studies on semiconductor deposition the deposition parameters (e.g., tip-sample voltage U_{em} , emission current I_{em}) were chosen based mainly upon empirical considerations and interpretation of the measurements was hampered by a lack of understanding of the (dominant) decomposition and deposition mechanisms.¹⁰ As a result, it was not easy to control the size of the deposited structures or to adopt the optimum deposition parameters to the specific precursor gas, partly because the exact processes during interaction of the precursor molecules with the electric field and the tip-emitted electrons in the region between the tip and the surface were not known. Several decomposition and deposition mechanisms have been discussed, such as field induced dissociation, dissociation at the tip, and subsequent field evaporation of the fragments, or dissociation by

interaction of the precursor molecules with the emitted electrons.¹⁰⁻¹²

We have tried to overcome this lack of understanding and to gain more information on the relevant decomposition process(es) because such knowledge is a necessary precondition for tuning the deposition parameters and optimizing the deposition process. Here we report results of a systematic analysis of the deposition behavior using both monosilane (MS) (SiH_4) and dichlorosilane (DCS) (SiH_2Cl_2) as precursors, which are also commonly employed for industrial silicon CVD. We are interested in particular in the deposition voltage threshold regime. Because of the limited space available here, we will restrict ourselves to the analysis of deposition voltage effects at positive sample bias and we will show that careful adjustment of this can lead to a substantial decrease of the dimensions of the deposited nanostructure as compared to that in our earlier work.¹⁰ Further, we can identify the process which, in the deposition voltage threshold regime, is the dominant one for decomposition and fragment accumulation on the surface for SiH_4 and probably also for SiH_2Cl_2 precursors.

II. EXPERIMENT

The experiments were carried out in an ultrahigh vacuum (UHV) system equipped with a pocket size STM, facilities for Auger electron spectroscopy, and a heating stage to flash the samples up to 1400 K by electron bombardment from the back. The samples were cut from *p*- $\text{Si}(111)$ wafers with a resistivity of 1–20 $\Omega\text{ cm}$. Preparation of the 7×7 reconstructed surfaces was carried out by following established procedures.²⁰ STM imaging was performed at sample voltages from -1.7 to -2.5 V. The images are displayed here in a gray scale with darker tones corresponding to lower areas in the topography. The gases are a high purity 2% $\text{SiH}_4/98\%$ Ar mixture or SiH_2Cl_2 that was additionally cleaned by a point of use Millipore Waferpure Micro system to reduce the impurities in the DCS to a few ppb.

Fabrication of the nanostructures via STM-CVD was carried out in a similar manner as that described in Ref. 10, namely, by exposing the surface to the precursor gas and

^{a)}Electronic mail: hubert.rauscher@chemie.uni-ulm.de

simultaneously applying a voltage of up to 25 V between the sample and the tip, with the current feedback loop on. In this mode the tip is in the field emission regime and free electrons emitted from the tip can interact with the precursor molecules; the latter may additionally be affected by the electric field between the tip and the sample. This way precursor is activated (decomposed or excited) and material can be deposited on the surface in the immediate vicinity of the tip by irreversible adsorption of the fragments or by dissociative adsorption of excited species. All deposition and subsequent imaging were performed at room temperature, where normal dissociative adsorption of MS and DCS precursor molecules saturates after submonolayer coverage.^{20,21}

III. SPOT FABRICATION AND THRESHOLD

The procedure described above results in the formation of surface nanostructures,¹⁰ provided the deposition voltage is above a certain threshold. By holding the tip at constant position and by careful optimization of the deposition parameters it is possible to deposit spotlike structures on the surface. Since the surface modifications caused by dissociative adsorption of the precursor gases SiH_4 or DCS with $\text{Si}(111)7\times 7$ at room temperature have been studied in detail earlier^{20,21} and are well known, STM induced depositions can be unambiguously identified on the surface.

This induced deposition is demonstrated in the STM image in Fig. 1, where a spot is displayed in a topview and a pseudo-three-dimensional representation. This spot was fabricated with the precursor DCS at a pressure of 5×10^{-6} mbar, a tip emission current of 38 pA, and the sample at a voltage of +12 V with respect to the tip, during a deposition time of 10 s. It represents one of the smallest spots fabricated on the surface by STM-CVD. A detailed inspection of the spot and its immediate vicinity on the surface reveals modifications in the 7×7 structure of the initial surface in the vicinity of the spot. The 7×7 structure is modified insofar as a fraction of the adatoms is darkened. The appearance of this part of the surface is identical to that of the $\text{Si}(111)7\times 7$ surface after interaction with DCS at room temperature,²¹ indicating that tip induced deposition is limited to a very narrow area with a diameter of about 150 Å in Fig. 1(a). The tip modified area can be classified into the spot itself, with a base width of 5.6 nm and a full width at half-maximum (FWHM) of 3.4 nm and a wider, speckled area with very small clusters or even additional adatoms compared to the initial surface. Both the spot and the latter features are clearly tip induced. The shape of the spot can be seen in more detail in the pseudo-three-dimensional (3D) representation of Fig. 1(b), where the scale in the vertical direction is enlarged. The shape of this spot closely resembles a regular cone with a height of 1.8 nm.

Another example of a spot fabricated at different parameters is displayed in Fig. 2, again in a pseudo-3D representation. This spot was deposited within 20 s at a sample voltage of +9 V, a current of 173 pA, and a pressure of $p(\text{DCS})=5 \times 10^{-6}$ mbar. This spot is directly deposited at a step of the initial surface, which, however, has no influence

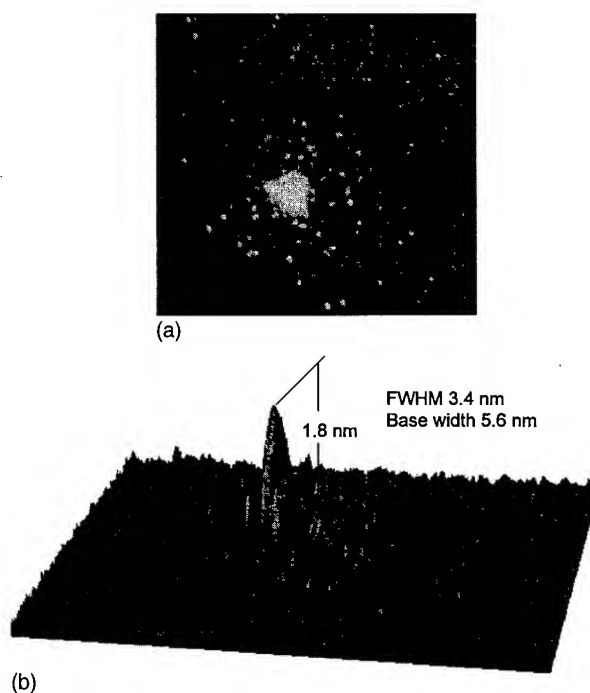


Fig. 1. (a) STM image of a spot-shaped nanostructure deposited from the precursor DCS on $\text{Si}(111)7\times 7$ at 300 K. In the vicinity of the spot the 7×7 surface after reaction with DCS is visible at atomic resolution. Spot dimensions: base width 5.6 nm, FWHM 3.4 nm, height 1.8 nm. Deposition parameters: $U_{em}=12$ V, $I_{em}=38$ pA, $p(\text{DCS})=5 \times 10^{-6}$ mbar, deposition time 10 s. Image size: 50 nm \times 47 nm. (b) Pseudo-3D representation of the spot in (a). The vertical scale is enlarged for better visualization of the spot shape.

on the deposited structure. Like in Fig. 1, the central spot is surrounded by a halo that can also be assigned to a tip induced modification of the surface. The dimensions of the spot in Fig. 2 (base width ~ 4 nm, FWHM 2.5 nm) are even

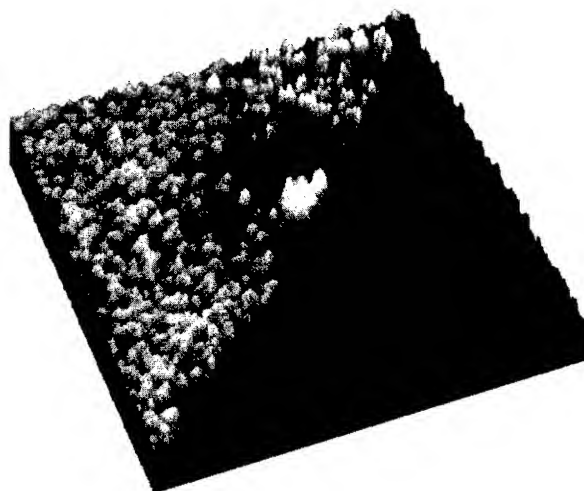


Fig. 2. Spot from DCS in pseudo-3D representation with the deposition voltage in the threshold regime. Spot dimensions: base width ~ 4 nm, FWHM ~ 2.5 nm, height 0.6 nm. Deposition parameters: $U_{em}=9$ V, $I_{em}=173$ pA, $p(\text{DCS})=5 \times 10^{-6}$ mbar, deposition time 20 s. Image size: 48 nm \times 49 nm.

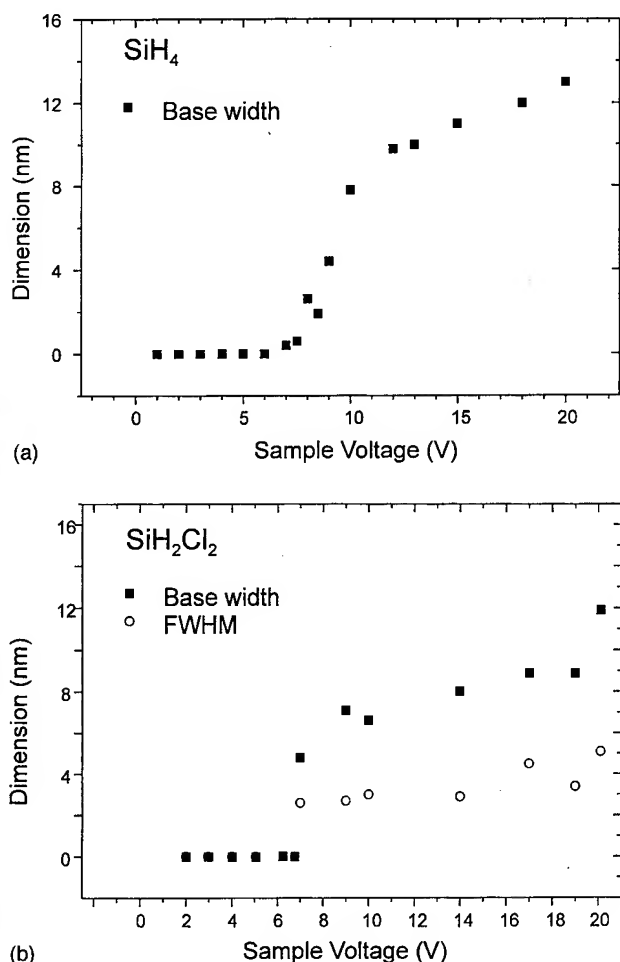


FIG. 3. Spot dimensions as a function of deposition voltage. (a) Precursor SiH_4 and (b) precursor SiH_2Cl_2 .

smaller than those of the spot in Fig. 1. In this second example the deposition voltage was close to the threshold where tip induced deposition of material from DCS is quite possible. These spots and the other structures presented here represent a substantial advance over our earlier results¹⁰ with respect to their dimensions and the ratio of their width to height.

As mentioned earlier we then performed a systematic study of the spot size as a function of deposition voltage utilizing the precursors monosilane and dichlorosilane. The results of these experiments are displayed in Fig. 3 for silane [Fig. 3(a)] and dichlorosilane [Fig. 3(b)] both at positive sample bias. The experiments with silane were carried out at a pressure of 2×10^{-6} mbar, a current of 38 pA, and a deposition time of 20 s; the parameters for the experiments with dichlorosilane were 5×10^{-6} mbar, 38 pA, and 20 s deposition time. The relevant dimension shown is the base width of the spots; for DCS [Fig. 3(b)] the FWHM is also included, which follows the base width.

Clear thresholds for deposition can be identified for the two precursors: 7.5–9 V for monosilane and 6.5–8 V for dichlorosilane. Voltages below 6 V never resulted in the formation of a spotlike deposition for either monosilane or

dichlorosilane. For both precursor gases the lateral dimensions and also the deposited number of species increase monotonously at voltages above the threshold.

IV. DISSOCIATION MECHANISM

Several deposition mechanisms have been considered in the literature to be effective for STM induced nanostructures using precursor gases. These include molecular decomposition on the tip and subsequent field evaporation, field dissociation and direct bond breaking, or a combination of them. To succeed in the reproducible fabrication of extremely small structures, however, precise knowledge of the dominant deposition mechanism(s) is essential in order to adjust the deposition parameters to specific systems to get the desired results.

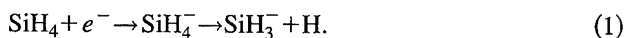
The threshold behavior observed in the voltage dependent deposition experiments (Fig. 3) leads to important conclusions regarding the relevant deposition mechanisms. In the low voltage regime where the experiments described above were conducted, several deposition mechanisms may be responsible for deposition in the threshold regime around 7–9 V. We first consider the precursor monosilane. Conceivable deposition mechanisms are (i) field induced dissociation, (ii) adsorption and dissociation at the tip and subsequent field evaporation onto the surface, (iii) direct ionization by electron impact, or (iv) interaction of the molecule (at the surface or still in the gas phase) with electrons by forming a temporary, unstable state and subsequent dissociation of the compound state. These mechanisms are discussed in the following.

- (i) Information on the exact electric field distribution and strength is difficult to obtain because the field in the immediate vicinity of the tip is very inhomogeneous and sensitive to the exact tip shape, which is unknown. We therefore estimated the electric field strength by modeling the tip by a metal sphere and the surface by a conducting, flat, structureless plane. The electric field resulting from such an arrangement can easily be calculated.²² The result is that in the threshold region the maximum electric field (at the tip) is in the range 0.1–0.5 V/Å, using a tip radius of 100 nm and a tip–surface distance of 15–50 Å. We would like to note here that these field strengths are too small to efficiently release surface atoms, a process that is well established,¹ but for which field strengths of the order of 4 V/Å and above are required.²³ Consequently, it was not possible for us to produce structures without precursor gas in the threshold regime similar to those presented above, but under otherwise identical parameters as those for the deposition experiments. On the other hand, field strengths in the range 0.1–0.5 V/Å appear to also be too small for direct field induced dissociation of silane, either adsorbed or in the gas phase,²³ so this mechanism is not suggested as being responsible for deposition in the threshold regime.
- (ii) Regarding the second possibility, the adsorption and

dissociation of molecules at the tip, it was reported earlier that voltages in the range of 1 V/Å are necessary for fragmentation of SiH₄ at the W tip at room temperature.¹¹ Subsequent processes that lead to material deposition would then be field-enhanced surface diffusion of the fragments to the tip apex and field evaporation to the Si surface. This sequence was favored in Ref. 11 but in our case the electric field strengths at the deposition threshold appear to be too small for this mechanism to become effective.

- (iii) The lowest ionization potential (IP1) for direct ionization of SiH₄ is at ~12 eV (ionization from the $2t_2$ orbital),²⁴ followed by ~18 eV for the $3a_1$ orbital (IP2).^{24,25} These values are clearly too high for the observed deposition threshold, so direct ionization can definitely be excluded. In addition, direct ionization would produce SiH₄⁺ ions which would be repelled from the surface at positive bias.
- (iv) It is well established that temporary negative ions (TNIs) can be formed by low energy electron capture of molecules.²⁶ When the electron energy matches the energy of a low-lying, unoccupied molecular orbital, the cross section for electron capture can increase considerably and a TNI resonance is formed. The TNI resonance can decay by dissociation under formation of an anion and a neutral radical. This process is called dissociative electron attachment (DEA). In fact, DEA was observed for SiH₄ earlier, with resonances at 7.0 and 8.0 eV, respectively.²⁷⁻²⁹ These resonance peaks for dissociative electron attachment were assigned to the Jahn–Teller split components of a valence TNI resonance involving a $t_2\sigma^*$ orbital.²⁹ This was confirmed by Dillon *et al.*³⁰ utilizing high resolution electron impact spectroscopy and by calculations by Tossell and Davenport,²⁵ which established that the t_2 resonance in SiH₄ arises from a single clearly identified unoccupied orbital.

The $t_2\sigma^*$ resonance transient anion state leads to molecular dissociation via DEA,



SiH₃[−] was identified as the most abundant species in the negative ion spectra of SiH₄ following low energy electron bombardment,²⁸ evolving via Eq. (1) with a cross-sectional maximum at ~7.0 eV. The threshold of 7.5–9 V for deposition from SiH₄ observed in our experiments is very close to the energies required for DEA in this molecule. We therefore suggest that this mechanism is the dominant one in the potential regime near the onset of deposition. Since deposition with the sample only at positive bias is considered here, the ions produced by DEA have the correct polarity to be attracted by the surface. This may also be the reason why deposition of material at the tip is not observed in the threshold regime. Only at higher voltages does tip contamination become more frequent. Results of experiments with negative sample bias, which will be discussed elsewhere, show that deposition at negative sample bias is considerably more

unreliable. To rationalize the slightly higher energy of the deposition threshold observed here compared to the gas phase TNI²⁹ one must keep in mind that under our conditions this process is strongly localized in space and restricted to a region above the surface where the electrons have gained enough energy necessary for the DEA process, which is smaller than the tip–surface distance. The necessary energy (i.e., ~7 eV) must have been gained some distance in front of the surface to enable the dissociation of a considerable number of molecules, which leads to significant material deposition.

DEA processes of adsorbed molecules have been well established and identified as very efficient surface dissociation processes, for example, for H₂O or physisorbed CF₄ on H-passivated Si(111) (Refs. 31 and 32) and O₂ on Si(111).³³ Because of the very specific excitation mechanism involved in DEA, it was realized that this process opens up the possibility of “tuning” chemical reactions at or near the surface by breaking specific chemical bonds.^{32,33} When using a STM tip as electron source for DEA, this process is strongly restricted in space, i.e., it is only possible in the region between the tip and the surface. This was utilized by Avouris and co-workers³⁴ for site and reaction specific STM induced surface chemistry. In our case, it was exploited for the fabrication of surface nanostructures. However, STM induced surface reactions of course do not necessarily need to be induced by DEA; other mechanisms specific to particular systems may also be effective for STM mediated surface modifications. This was explicitly shown for STM induced H desorption from Si(111) (Ref. 13) or Si(100).^{14,34}

For dichlorosilane the situation is less clear than for monosilane in that there are no equally reliable data for TNI and DEA processes for this molecule, but the processes and aspects discussed in (i)–(iii) (field dissociation, field evaporation, and direct ionization) may hold for this molecule as well. Direct ionization, in particular, can be excluded in the threshold regime, because the first IP for DCS is 11.7 eV.³⁵ Regarding the deposition threshold of 6.5–8 V, which is slightly lower than that for monosilane, a similar TNI resonance that leads to DEA may well exist, and we tentatively assign the threshold behavior for DCS to a DEA process similar to that for silane.

At potentials well above the threshold region, processes other than DEA will certainly contribute to the dissociation of SiH₄ and SiH₂Cl₂; these were discussed previously. With increasing deposition voltage processes like direct ionization or field induced dissociation will also become effective and lead to less specific behavior in the deposition. This will lead also to a less pronounced localization of the dissociation process and the molecular fragments on the surface and will result in larger dimensions of the nanostructures.¹⁰ Therefore, to achieve the minimum size for nanostructures it is useful to deposit material with voltages as close to the deposition threshold as possible; this, on the other hand, results in a relatively slow “writing” speed, i.e., the time necessary to produce a structure of the desired size.

We would like to add a remark regarding the chemical

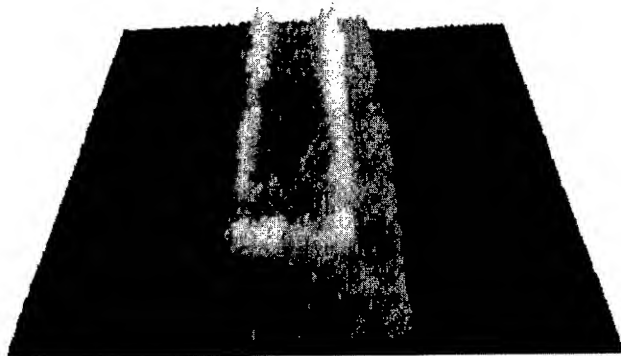


FIG. 4. Pseudo-3D representation of a continuous structure written from the precursor DCS. Structure dimensions: linewidth 9 nm, distance between the parallel lines 40 nm, line height 1 nm. Deposition parameters: $U_{em} = 12$ V, $I_{em} = 32$ pA, $p(\text{DCS}) = 5 \times 10^{-5}$ mbar, writing speed 10 nm/s. Image size: 208 nm \times 120 nm.

composition of the structures. Since the experiments were carried out far below the desorption temperature of H_2 from Si, we expect that the deposits are composed of amorphous, hydrogenated and, in the case of deposition from DCS, chlorinated Si. This is also suggested by the dissociation mechanism(s) which produce(s) a mixture of SiH_x or SiH_xCl_y , plus H species. $(dI/dV)/(I/V)$ measurements (not displayed here) over the nanostructures indicate that the characteristic surface states observed on bare Si(111) 7×7 , which result from the surface dangling bonds, are not present on these structures. This suggests saturation of the Si dangling bonds, which can be expected in hydrogenated/chlorinated Si clusters.

The information gained from spot deposition experiments and from analysis of the deposition parameters in the threshold region can be used to fabricate continuous and rather narrow surface structures by moving the tip under deposition conditions. An example is shown in Fig. 4 in a pseudo-3D representation. This structure was deposited close to a step, which can be seen at the right-hand side of the image. The precursor was DCS at a pressure of 5×10^{-5} mbar, $U_{em} = 12$ V, $I_{em} = 32$ pA, at a writing speed of 10 nm/s. The deposited structure has a regular U shape, without any interruptions. The linewidth in this structure is 9 nm, and the parallel lines are separated by ~ 40 nm. Continuous structures having such narrow linewidths could be fabricated only at voltages close to the threshold regime, confirming that information on the deposition threshold regime is important for the reproducible fabrication of nanostructures.

V. SUMMARY

Nanostructure fabrication by STM-CVD was studied on Si(111) 7×7 surfaces utilizing the precursor gases silane (SiH_4) and dichlorosilane (SiH_2Cl_2), with emphasis on the behavior in the voltage threshold regime. Spotlike as well as continuous structures with lateral dimensions in the sub-10 nm regime were successfully fabricated by tuning the depo-

sition parameters. The dissociation behavior strongly suggests that for SiH_4 decomposition a dissociative attachment process becomes effective in the deposition voltage threshold regime, and may also apply to DCS as well.

ACKNOWLEDGMENT

The authors thank the German Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF) for financial support through Grant No. 13N6322A. The silicon samples were kindly provided by Wacker Chemitronic.

¹*Atomic and Nanometerscale Modifications of Materials: Fundamentals and Applications*, edited by Ph. Avouris (Kluwer, Dordrecht, 1993).

²R. Wiesendanger, J. Vac. Sci. Technol. B **12**, 515 (1994).

³R. M. Silver, E. E. Ehrichs, and A. L. de Lozanne, Appl. Phys. Lett. **51**, 247 (1987).

⁴E. E. Ehrichs, S. Yoon, and A. L. de Lozanne, Appl. Phys. Lett. **53**, 2287 (1988).

⁵M. A. McCord and D. D. Awschalom, Appl. Phys. Lett. **57**, 2153 (1990).

⁶S. T. Yau, D. Saltz, and M. H. Nayfeh, J. Vac. Sci. Technol. B **9**, 1371 (1991).

⁷E. E. Ehrichs, W. F. Smith, and A. L. de Lozanne, Ultramicroscopy **42-44**, 1438 (1992).

⁸S. Rubel, M. Trochet, E. E. Ehrichs, W. F. Smith, and A. L. de Lozanne, J. Vac. Sci. Technol. B **12**, 1894 (1994).

⁹S. Rubel, X.-D. Wang, and A. L. de Lozanne, J. Vac. Sci. Technol. B **13**, 1332 (1995).

¹⁰H. Rauscher, U. Memmert, and R. J. Behm, J. Vac. Sci. Technol. B **13**, 1216 (1995).

¹¹T. M. H. Wong, S. J. O'Shea, A. W. McKinnon, and M. E. Welland, Appl. Phys. Lett. **67**, 786 (1995).

¹²D. Samara, J. R. Williamson, C. K. Shih, and S. K. Banerjee, J. Vac. Sci. Technol. B **14**, 1344 (1996).

¹³R. S. Becker, G. S. Higashi, Y. J. Chabal, and A. J. Becker, Phys. Rev. Lett. **65**, 1917 (1990).

¹⁴J. W. Lyding, T.-C. Shen, J. S. Hubacek, J. R. Tucker, and G. C. Abeln, Appl. Phys. Lett. **64**, 2010 (1994).

¹⁵M. A. McCord and R. F. W. Pease, J. Vac. Sci. Technol. B **5**, 430 (1987).

¹⁶F. K. Perkins, E. A. Dobisz, S. L. Brandow, J. M. Calvert, J. E. Kosakowski, and C. R. K. Marrian, Appl. Phys. Lett. **68**, 550 (1996).

¹⁷D. M. Eigler and E. K. Schweizer, Nature (London) **344**, 524 (1990).

¹⁸H. Uchida, D. Huang, F. Grey, and M. Aono, Phys. Rev. Lett. **70**, 2040 (1993).

¹⁹G. Dujardin, R. E. Walkup, and Ph. Avouris, Science **255**, 1232 (1992).

²⁰U. Memmert, A. Berkó, and R. J. Behm, Surf. Sci. **325**, L441 (1995).

²¹M. Fehrenbacher, H. Rauscher, and R. J. Behm, Z. Phys. Chem. **198**, 205 (1997); **199**, 123 (1997).

²²M. A. McCord and R. F. W. Pease, J. Vac. Sci. Technol. B **3**, 198 (1985).

²³H. J. Kreuzer, in *Chemistry and Physics of Solid Surfaces VII*, edited by R. Vanselow and R. Howe (Springer, Berlin, 1990).

²⁴H. Bock, W. Ensslin, F. Feher, and R. Freund, J. Am. Chem. Soc. **98**, 668 (1976).

²⁵J. A. Tossell and J. W. Davenport, J. Chem. Phys. **80**, 813 (1984).

²⁶L. Sanche, Scanning Microsc. **9**, 619 (1995).

²⁷H. Ebinghaus, K. Kraus, W. Müller-Duysing, and H. Neuert, Z. Naturforsch. A **19**, 732 (1964).

²⁸P. Potzinger and F. W. Lampe, J. Phys. Chem. **73**, 3912 (1969).

²⁹M. B. Robin, *Higher Excited States of Polyatomic Molecules* (Academic, Orlando, 1985), Vol. 3.

³⁰M. A. Dillon, R.-G. Wang, Z.-W. Wang, and D. Spence, J. Chem. Phys. **82**, 2909 (1985).

³¹W. Di, P. Rowntree, and L. Sanche, Phys. Rev. B **52**, 16618 (1995).

³²D. Klyachko, P. Rowntree, and L. Sanche, Surf. Sci. **346**, L49 (1996).

³³R. Martel, Ph. Avouris, and I.-W. Lyo, Science **272**, 385 (1996).

³⁴Ph. Avouris, R. E. Walkup, A. R. Rossi, H. C. Akpati, P. Nordlander, T.-C. Shen, G. C. Abeln, and J. W. Lyding, Surf. Sci. **363**, 368 (1996).

³⁵D. C. Frost, F. G. Herring, A. Katrib, R. A. N. McLean, J. E. Drake, and N. P. C. Westwood, Chem. Phys. Lett. **10**, 347 (1971).

Studies of field related effects in the fabrication process on graphite using a scanning tunneling microscope

Chen Wang, Xiaodong Li,^{a)} Guangyi Shang, Xiaohui Qiu, and Chunli Bai
Institute of Chemistry, Chinese Academy of Sciences, Beijing 100080, People's Republic of China

(Received 6 March 1997; accepted 10 March 1997)

The surface modification process with low energy electron beams, generated with a scanning tunneling microscope, is investigated using graphite as an example. By analyzing the cross-sectional profiles, it is shown that local transport properties could play a vital role in determining the characteristics of the fabricated craters. In addition, the observed asymmetry in the contours of the as-produced craters is proposed to be correlated with the geometry of the tip apex region. It is therefore suggested that one look into the roles of low energy electrons, together with the effects of strong electric field, in the fabrication process. © 1997 American Vacuum Society. [S0734-211X(97)01204-3]

I. INTRODUCTION

As considerable progress has been made concerning the physical nature of manipulation and the fabrication process using scanning tunneling microscopy (STM),^{1,2} it is becoming increasingly clear that the fabrication and controlling processes are complicated issues and require knowledge of not only individual atoms but also species of condensed form as well. Currently a consensus has not been fully reached on the fabrication mechanisms. It is evident that the fabrication process using STM is affected by many factors, some of which are strong electric field and injected low energy electrons—two of the most common ones in this category of experiments, and their importance has been widely recognized. While the roles of strong electric field, along with other proximity effects, have been extensively studied, the possible effects of the injected electrons have yet to be clearly understood. The interaction of low energy electrons with condensed matter and molecular solids has been an interesting subject over the years,³⁻⁵ and is known to be associated with a number of phenomena. As progress is seen in revealing the mechanisms for atom manipulation, it is naturally interesting to explore the range of parameters where the mechanisms could be applied.

Recent theoretical^{6,7} and experimental⁸ studies strongly support the view that the injected electron beams are well collimated within a narrow range of nanometer scale. The apparently large sizes of modified regions on the surface of a number of materials such as graphite,^{9,10} TaS₂,¹¹ and so on indicate that there should be a diffusion mechanism involved.

Our previous study of the diffusion process for low energy electrons¹² suggested that the cross-sectional profiles of craters generated by electric pulses could be influenced by the diffusion of injected electrons due to various scattering processes. Comparisons with experimental data on highly oriented pyrolytic graphite (HOPG) yielded qualitatively satisfactory results.¹² It is further suggested that the correlation

could be applied to study the local transport properties of sample materials. While this consideration was mainly based on the analysis of highly symmetric cross-sectional profiles, we also noticed that there are occasions where substantial asymmetric results exist, which were also apparently dependent on the tips used in the experiments. It is plausible to assume that the tip geometry should have certain impacts on the fabricated features, since it is directly related to the distribution of the electric field external to the sample surface. Therefore, tip geometry is an issue that should be addressed while analyzing the characteristics of the concerned nanoscale features. In subsequent investigations we also demonstrated that two distinct types of features can be repeatedly produced and are found to be correlated with tip-sample separation, displaying threshold behavior.¹³ This observation could be considered as consistent with the general trend of tunneling barrier versus tip-sample separation.

In this work, we examined the characteristics of the craters in more detail. Evidence of asymmetric cross-sectional profiles and associated contour plots was presented, and the explicit expression of correlation between the geometry of tip apex region and the distribution of electric field is derived. The combination of the electron diffusion process and the electric field distribution, which is external to the sample surface, helped to illustrate the roles of field associated factors. We analyzed the possibility of obtaining information on local transport properties and geometry of tip apex by examining the characteristics of cross-sectional profile of as-produced craters.

II. CHARACTERISTICS OF THE ELECTRON DIFFUSION PROCESS

There has been evidence of effects due to high current density for many years.^{14,15} Attention has been directed at the possible Joule heating effects, as this correlation is believed to be minimal. So far, research has been focused on the effects of strong electric field, and there is apparently a lack of discussion on the possible interactions between the accompanying low energy electrons with atoms of the sample during the fabrication process, although this interac-

^{a)}Permanent address: Dept. of Materials Physics, University of Science and Technology of Beijing, Beijing 100083, People's Republic of China.

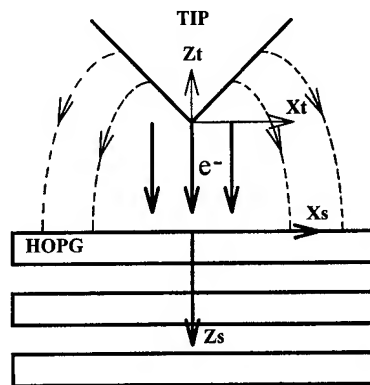


FIG. 1. A schematic diagram of tip-sample configuration illustrating the fabrication process using STM.

tion has been known to be the origin of destabilizing target atoms, which is also a subject being extensively studied by atomic physicists.³⁻⁵

In an idealized situation, the injected electron beam in the fabrication operation using STM can be modeled as a point source. By positioning the incoming electron beam at the origin of the coordinates as shown in Fig. 1, one can illustrate the distribution of the electron in terms of a diffusion equation:¹²

$$D_{\parallel} \frac{\partial^2 \phi}{\partial x_s^2} + D_{\parallel} \frac{\partial^2 \phi}{\partial y_s^2} + D_{\perp} \frac{\partial^2 \phi}{\partial z_s^2} = -W(z_s) \delta(x_s, y_s).$$

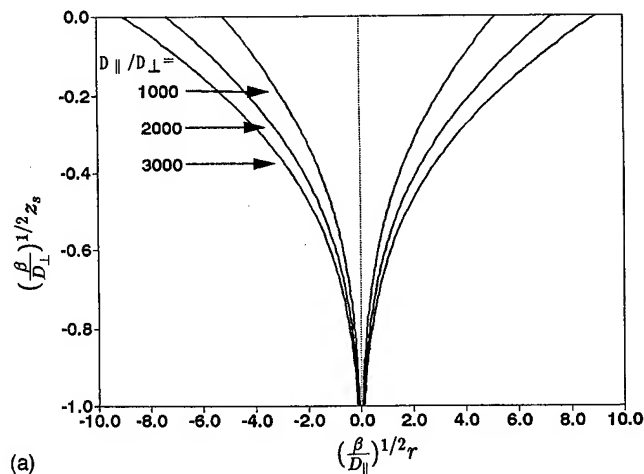
Here (x_t, z_t) and (x_s, y_s, z_s) represent tip and sample surface positions, respectively. y_s , which is not shown in the figure, is the axis normal to the paper plane. $W(z_s)$ is the number of electrons arriving per unit length along the path of the injected electrons per unit time. ϕ is the density of electrons in the vicinity of the injection position. D_{\parallel} and D_{\perp} are the diffusion constants parallel and perpendicular to the surface, respectively.

The equation is a typical one for Green's function and can be solved analytically. The general solution at the vicinity of the tip apex is in the form of

$$W(z_s) = W_0 \exp \left[- \left(\frac{\beta}{D_{\perp}} \right)^{1/2} z_s \right],$$

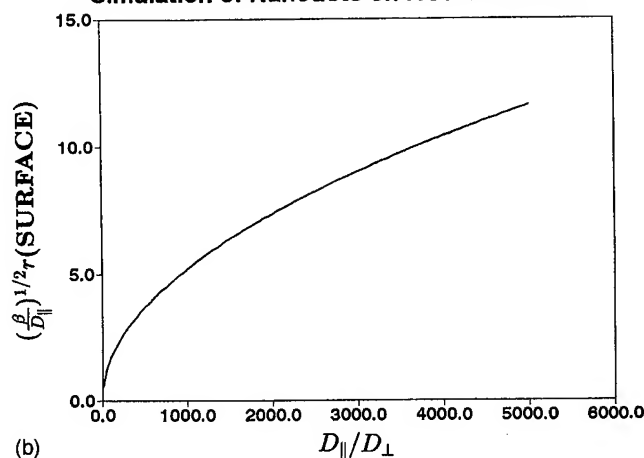
$$\phi = - \frac{W_0}{2\pi D_{\parallel}} \ln \left[\left(\frac{\beta}{D_{\parallel}} \right)^{1/2} \frac{r}{2} \right] \exp \left[- \left(\frac{\beta}{D_{\perp}} \right)^{1/2} z_s \right].$$

Here W_0 and β are constants. $r = \sqrt{x_s^2 + y_s^2}$. The exponential decay of beam intensity is evident in the above expression. Since the density distribution is closely related to the diffusion properties, it would be interesting to see the effects caused by the variation of D_{\parallel} and D_{\perp} . Figure 2(a) displays an example of the calculational results of constant density profile for three different D_{\parallel}/D_{\perp} ratios, where a value of 3000 is derived from the parameters of graphite.^{12,16} The effect of D_{\parallel}/D_{\perp} on the surface diameter of the affected region is clearly visible, and is also presented in Fig. 2(b). The negative z range represents the sub-surface region. The result indicates that materials having a higher ratio of in-plane to



(a)

Simulation of Nanodots on HOPG Surface



(b)

FIG. 2. (a) Different D_{\parallel}/D_{\perp} ratios will result in different constant density profiles; (b) the variation of surface radius as the function of D_{\parallel}/D_{\perp} .

out-of plane diffusion constants will be more susceptible to extended surface damages during the nanostructuring process involving field emitted electrons as the major fabrication tool. As to the experimental test of this analysis, a model system would be the graphite intercalation compounds which are layered systems with various asymmetries of transport properties,¹⁶ which therefore could serve as ideal candidates for this purpose.

The above analysis mainly used an idealized scheme where a point source is the origin of the field emission electrons. It could easily be extended to the situation of multiple emission. The multiple sources would then lead to the superimposed distribution of electrons. This is evidenced by some observations that the lowest portion of some generated craters are not pointlike, but rather a flat region of nanometer sizes. We believe this is an indication of a relatively flat tip. Another possibility for generating the flat region is that the as-generated craters could undergo a collapsing process due to structural instabilities or migrating surface atoms.¹² In view of the fact that a relatively long time is required for such a process to occur, we believe this is not the dominant mechanism, and therefore could be ruled out.

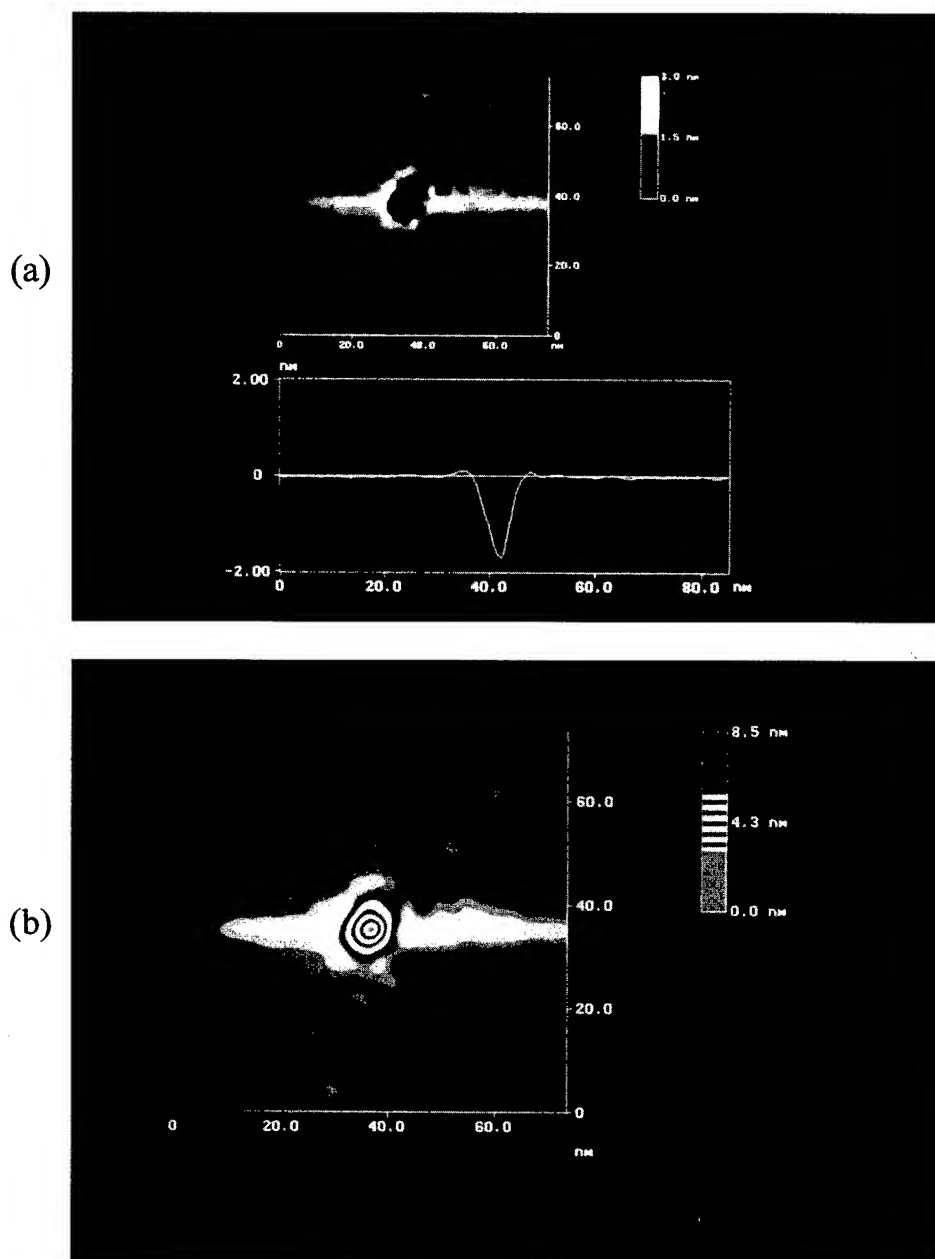


FIG. 3. STM image of a nanometer sized features on the surface of HOPG obtained by applying a 8 V, 3 ms voltage pulse to the sample at a set point of (500 mV, 50 pA). (a) The cross-sectional profile showing the crater is relatively high symmetric; (b) corresponding contour image.

III. TIP GEOMETRY EFFECT

In addition to the observation of deviation from an idealized situation of point emission source as discussed in Sec. II we also noticed asymmetries in overall contours of certain nanoscale craters. Numerous investigations have shed light on the roles of a strong electric field in fabrication processes. Mechanisms such as field initiated contact,¹⁷ field evaporation,¹⁸ and so on presented different approaches to the situation. As a complement to the above mentioned approaches, it would be interesting to look at the actual distribution of the electric field. By taking into account the charge distribution on the tip surface near its apex region, one could

obtain an explicit form of the electric field distribution E (directed normal to the sample surface) at the sample surface which is presumably grounded:

$$E(x_s) = 2k \int_{\text{tip}} \frac{z_t \sigma(x_i, z_t)}{[z_t^2 + (x_s - x_i)^2]^2} \left[1 + \left(\frac{dz_t}{dx_t} \right)^2 \right]^{1/2} dx_t.$$

Here the coordinates are defined as in Fig. 1. $E(x_s)$ is the electric field at surface position x_s , $\sigma(x_i, z_t)$ is surface charge density at position (x_i, z_t) on the tip, and dz_t/dx_t is the local curvature of the tip. We use two dimensional coordinates as shown in Fig. 1 for the convenience of illustration. An extension to the three dimensions could be readily avail-

able. For simplicity, one could consider spherical, parabolic, and hyperbolic models, respectively, for the tip apex in order to obtain $\sigma(x_t, z_t)$ explicitly, similar to the treatments widely used in field emission studies. In practice, the tip shape may very likely involve a more complicated geometry.

The above expression clearly involves geometric information on the tip. Since the field strength is concentrated at the tip apex region, the field emission in the case of STM is reasonably considered to be restricted to the tip apex, especially when the pulse voltage is not much larger than the threshold value. In addition, the field strength outside the field emission region may still be strong enough to affect the fabrication process, causing variations in the apparent geometry of the fabricated features. This analysis implies that it might be possible to retrieve information about actual tip apex geometry from fabrication outcomes.

IV. EXPERIMENTAL OBSERVATIONS

Experimental observations of the fabricated craters on HOPG surface have yielded qualitatively consistent results within the framework of the above considerations. Cross sections of craters with relatively high symmetry, as shown in Fig. 3(a) could be closely simulated by invoking the electron diffusion mechanism.¹² This feature was obtained with the tunneling conditions (500 mV, 50 pA) and with a voltage pulse of (−8 V, 3 ms) applied to the tip. The feedback was turned off in order to stabilize the tip-to-sample separation. The corresponding contour image in Fig. 3(b) helps to illustrate the symmetry of the crater. For some of the craters which have limited diameters at the bottom, yet are still symmetric, reasonable fits can still be obtained after removing the flat bottom region. This helps to show the effects of flat tips. We also have examples where two adjacent craters were generated simultaneously by the pulse method. This further suggests that the emission source could consist of multiple points. It was noticed that in the fabrication process on a HOPG surface, a nonsymmetric configuration could also be observed. This is a clear indication that there are other factors that could affect the fabrication process. As discussed in Sec. III variations in the crater symmetry could be accounted for by considering the effects of tip geometry.

Tip effects have been studied with an emphasis on the proximity effect, which causes considerable deformation of local tunnel barrier, thus increasing the probability for the transfer of atoms between the tip and the sample. This is highly convincing in the situation of atom manipulation.

However, for fabrication of regions as large as tens of nanometers, proximity effect alone appear not to be sufficient. Other mechanisms, including chemical ones, probably should be taken into account in order to provide more specific correlations between various electric field related factors. This may facilitate quantitative descriptions of the fabrication process employing STM.

V. SUMMARY

This investigation suggests that the characteristics of the nanometer scale craters could be appreciably influenced by the local transport properties, and tip geometry as well. It therefore provides a possible venue to retrieve relevant information which is not readily available by other means. Further efforts are considered necessary and essential in order to gain a deeper insight into the fabrication process using STM.

ACKNOWLEDGMENTS

The work is supported by the Foundation of Chinese Academy of Sciences and the National Natural Science Foundation of China. One of the authors (C.W.) is grateful for the support from the Director's Fellowship of the Institute of Chemistry, CAS.

¹J. W. Lyo and P. Avouris, *Science* **253**, 173 (1991).

²J. A. Stroscio and D. M. Eigler, *Science* **254**, 1319 (1991).

³G. J. Schulz, *Rev. Mod. Phys.* **45**, 378 (1973).

⁴L. Sanche, *Scanning Microsc.* **9**, 619 (1995).

⁵U. Fano and J. A. Stephens, *Phys. Rev. B* **34**, 438 (1986); U. Fano, *ibid.* **36**, 1929 (1987).

⁶K. Hirose and M. Tsukada, *Phys. Rev. B* **51**, 5278 (1995).

⁷B. Das and J. Mahanty, *Phys. Rev. B* **36**, 898 (1987).

⁸V. T. Binh, S. T. Purcell, N. Garcia, and J. Doglioni, *Phys. Rev. Lett.* **69**, 2527 (1992).

⁹T. R. Albrecht, M. M. Dovek, M. D. Kirk, C. A. Lang, C. F. Quate, and D. P. E. Smith, *Appl. Phys. Lett.* **55**, 1727 (1989).

¹⁰Wang Zhonghuai, Dai Changchun, Zhang Pingcheng, Huang Guizhen, Li Renli, Guo Yi, and Bai Chunli, *Chin. Phys. Lett.* **10**, 535 (1993).

¹¹Chen Wang, Zhi Zhao, and Chunli Bai (to be published).

¹²C. Wang, C. L. Bai, X. D. Li, G. Y. Shang, I. Lee, X. W. Wang, X. H. Qiu, and T. Fang, *Appl. Phys. Lett.* **69**, 348 (1996).

¹³C. Wang, X. D. Li, G. Y. Shang, X. H. Qiu, and C. L. Bai, *J. Appl. Phys.* **81**, 1227 (1997).

¹⁴Y. Z. Li, L. Vazquez, R. Piner, R. P. Andres, and R. Reifenberger, *Appl. Phys. Lett.* **54**, 1424 (1989); P. F. Marella and R. F. Pease, *ibid.* **55**, 2366 (1989).

¹⁵K. Koning, O. Jusko, L. Koenders, and A. Schlachetzki, *J. Vac. Sci. Technol. B* **14**, 48 (1996).

¹⁶M. S. Dresselhaus and G. Dresselhaus, *Adv. Phys.* **30**, 139 (1981).

¹⁷J. I. Pascual, J. Mendez, J. Gomez-Herrero, A. M. Baro, N. Garcia, and V. T. Binh, *Phys. Rev. Lett.* **71**, 1852 (1993).

¹⁸T. T. Tsong, *Phys. Rev. B* **44**, 13703 (1991).

Nanostructuring by reactive accelerated cluster erosion

A. Gruber and J. Gspann

Universität Karlsruhe und Forschungszentrum Karlsruhe, Institut für Mikrostrukturtechnik, Karlsruhe, Germany

P. von Blanckenhagen

Forschungszentrum Karlsruhe, Institut für Materialforschung I, Karlsruhe, Germany

(Received 6 March 1997; accepted 10 March 1997)

Accelerated ionized cluster beams are used for surface nanostructuring of bulk diamond, chemically vapor deposited (CVD) diamond films, single-crystalline silicon, or glass, among others. Beams of CO₂ clusters with about 1000 molecules per unit charge are accelerated to up to 120 keV kinetic energy for mask projective surface bombardment. Patterning is achieved via physical as well as chemical surface erosion. Very smooth eroded surfaces result for bulk natural diamond, silicon, and glass. Polycrystalline, strongly faceted CVD diamond films are effectively planarized. Submicrometer structures with adjustable wall inclinations can be generated. Atomic force microscopy of isolated impact structures reveals hillocks instead of craters. The collective motion of the impacted surface material is considered crucial for the cluster impact induced nanomodifications. © 1997 American Vacuum Society. [S0734-211X(97)03104-1]

I. INTRODUCTION

Beams of ionized clusters can be used for impact induced micromachining.¹ Electrical acceleration by 100 kV of ionized clusters with 1000 molecules per charge unit leads to physical² as well as chemical³ impact erosion. The high temperatures that are calculated to result at the impact area from the high impact energies^{4,5} provide etch reactivity, at least for some ps, even with otherwise rather inert cluster materials, such as CO₂ or SF₆. Hence, the process is called reactive accelerated cluster erosion (RACE).³

The present study is devoted to investigating the nanostructures generated by individual cluster impacts. The lateral extensions of these nanostructures should determine the limits of the lateral resolution that is achievable with a cluster beam lithography, while the depths or heights of the nanostructures should add up to the roughness of the eroded surface.

II. EXPERIMENT

The clusters are generated by adiabatic nozzle expansion of CO₂ or SF₆ gas, the converging-diverging nozzle having 0.1 mm throat diameter, a 10° angle of initial divergence, and a 28 mm long divergent part (Fig. 1). Most of the expanding gas is frozen onto cryopanel attached to a liquid nitrogen bath cryostat. The core of the expanding nozzle flow is transferred to high vacuum via two skimming orifices. Electrons of 150 eV energy partly ionize the cluster beam which may then be focused by up to a 10 kV potential negative with respect to the acceleration potential.⁶

The cluster mean sizes can be measured by using a dedicated time-of-flight spectrometer in place of the target.⁷ The nozzle feed pressure is chosen to provide clusters of a suitable mean size (in the present work 1000 molecules of CO₂ per unit charge) in order to ensure high speed impacts. At a chosen acceleration voltage of 100 kV the ionized clusters impinge on the target at about a 20 km/s speed.

The experimental setup described in Ref. 6 earlier is now equipped with a rapid beam shutter that allows exposure times as short as 0.5 ms. A manually operated beam flag additionally serves to protect the rapid shutter from prolonged cluster beam erosion. Target, mask, and beam shutters are all kept at ground potential. The cluster beam source can be operated at high potential of either polarity. In the present case, a positive potential of 100 kV is used to accelerate positive cluster ions towards the ground electrode which they pass via a 10-mm-diam central orifice.

After exposure to the accelerated ionized cluster beam, the targets are transferred through ambient atmosphere without any particular precautions to either a scanning electron microscope or a Digital Instruments Nanoscope III. They are investigated by atomic force microscopy in the so-called contact mode using silicon nitride cantilevers, or in the tapping mode with Si cantilevers.

III. RESULTS

Direct lithography by accelerated cluster beams is exemplified in Fig. 2 by the erosion and planarization of chemically vapor deposited (CVD) diamond films. The strongly faceted films typical of CVD diamond are eroded using a mask with hexagonal holes. Originally, the films were black due to a residual graphitic component, which also makes the films conducting. In the scanning electron microscope (SEM), these parts appear bright. The eroded parts, however, appear black in the SEM due to lost conductivity, but appear nearly white under natural illumination.

The planarization of the eroded areas gives the impression of superficial melting. The underlying crystal structure can still be recognized in the residual, mostly triangular holes.

With homogeneous targets, such as monocrystalline silicon, metals, glass, and others, the eroded areas have always been found to be very smooth,^{1,3,6} at least as seen in SEM

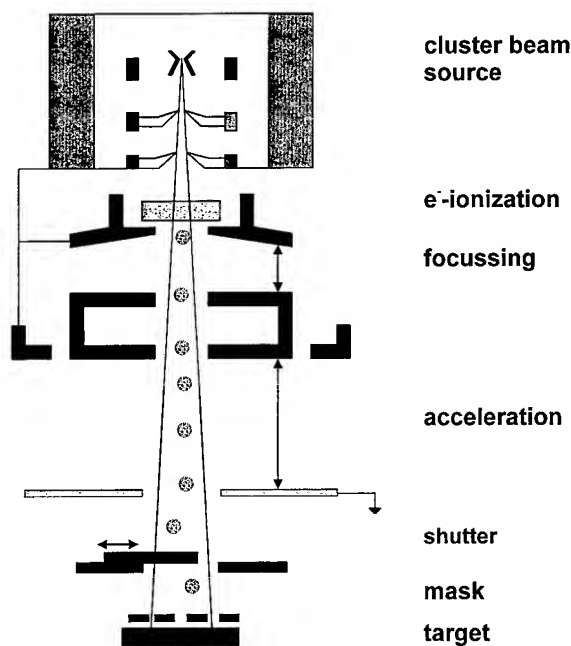


FIG. 1. Schematic view of the setup for generating isolated cluster impact structures.

investigations. In the following, results obtained in atomic force microscope (AFM) studies are presented; they are used to clarify the individual impact structures.

Commercially available polished Si wafers were selected as targets because of their original flatness. Figure 3 gives an example of such a surface after CO_2 cluster bombardment for 0.125 s investigated by AFM in the contact mode. Surprisingly, there appears to be a rise in the eroded area instead of a depression. In Figs. 3 and 4 light shading indicates elevated regions.

In Fig. 4, enhanced spatial resolution is used to envision a typical isolated hillock. The areal density of such hillocks corresponds to the areal density of cluster impacts. It was originally expected that the result of an individual cluster impact would be a nanosized crater.²

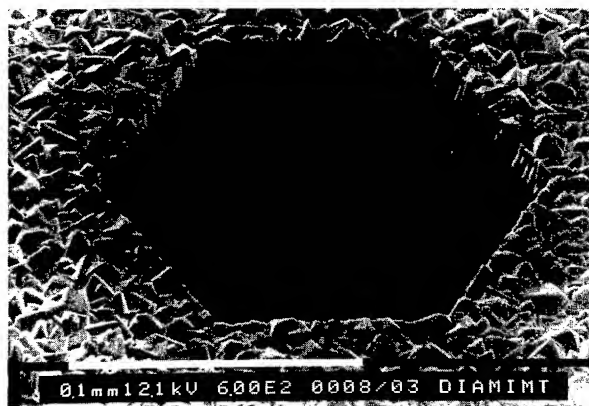


FIG. 2. Polycrystalline diamond eroded by accelerated CO_2 cluster bombardment. The hexagonal eroded area appears black in the SEM since a loss of residual graphite reduces the electrical conductivity.

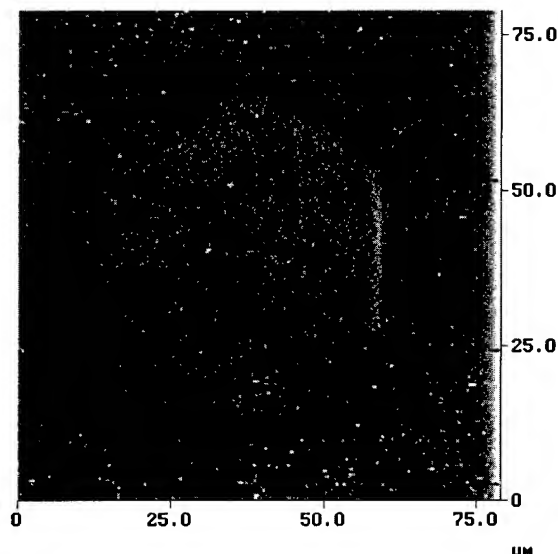


FIG. 3. Silicon after 0.125 s bombardment with accelerated CO_2 clusters. The AFM picture is a deflection image taken in the contact mode. The impacted area seems to be raised due to the image of many hillocks combined.

Figure 5 gives an example of a much larger impact structure. Such patterns occur about a factor of 1000 less frequently than the isolated hillocks in Figs. 3 and 4 and must result from much larger and more energetic projectiles. The origin of these larger projectiles is not quite clear. Presumably they are multiply charged, allowing them to gain a correspondingly higher energy in the electrostatic acceleration. The central pyramid in Fig. 5 is surrounded by an area that again indicates the flow of molten material.

IV. DISCUSSION

The spatial resolution of direct microlithography by cluster impact erosion must be limited by the spatial extension of the material modification induced by an individual impact. From macroscopic ballistic experiments it is known that hypervelocity impacts of spherical projectiles generally yield hemispherical surface craters whose volume is proportional to the impact energy. Extrapolating these results to the im-

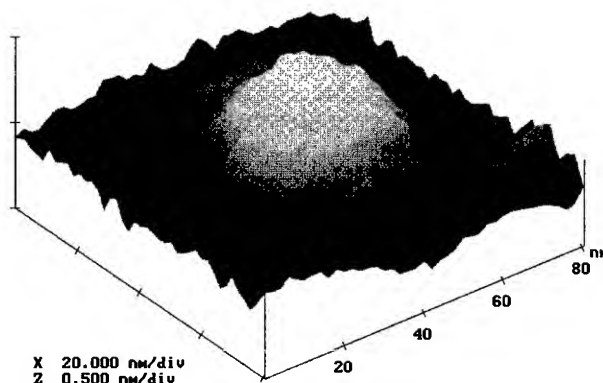


FIG. 4. Isolated impact-induced hillock on silicon, observed as the height image in the tapping mode.

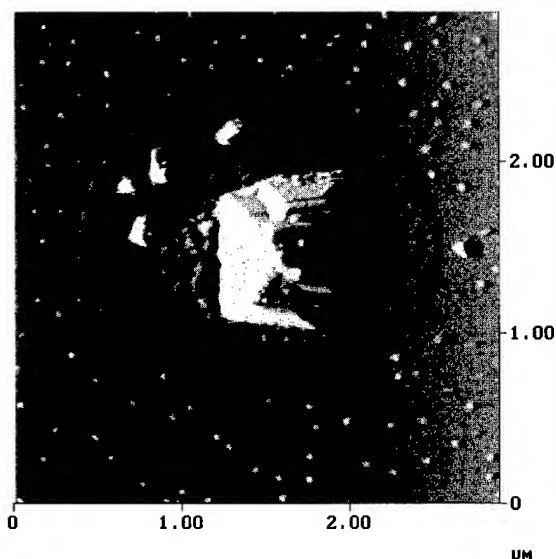


FIG. 5. Impact-induced structure on silicon with a central pyramid and surrounding flow blanket. The pyramid is considered to be a convolution of a central peak on the target by the pyramidal contact mode AFM tip.

pacts of clusters of 100 keV impact energy leads to predicted crater diameters of some tens of nm, which is considerably larger than the average diameter of about 3 nm of the impinging clusters.² Indeed, a pronounced surface roughening has already been observed with atomic force microscopy after hypervelocity cluster bombardment.⁸ While the lateral dimensions of the hills or pits were of the order of some tens of nm, it was not possible, however, to decide whether the surface structure resulted from an interference by craters or by hills.

Figures 3 and 4 now reveal the isolated structures to be hillocks which have the approximate lateral dimensions anticipated. Hence, while the original considerations may still be used with respect to the lateral limit of resolution for this kind of lithography, the low height of the resulting hillocks helps to explain the observed smoothness of the eroded areas.

The accelerated clusters of 1000 molecules impinge onto the target at about a 20 km/s speed which is more than the

sound speed in any material. Hence, a shock wave will traverse the target radially from the point of cluster impact, setting the target material into a radially outward directed motion also. If a hemispherical crater does not result from such an impact, this outward motion must later be reversed. The rebounding material may then refill the transient crater, eventually even leading to a central peak. Such a process has also been postulated to explain the very large impact structures with a central peak on the planet Mars.⁹

The central pyramid in Fig. 4 is very probably a convolution of a central peak on the target and the tip of the AFM. If more than one such pattern can be found on the target, they should all be similar in orientation and topology. Furthermore, the inclinations of the sideplanes of the pyramids are the same as the inclinations obtained in the anisotropic etching of Si which is used in the tip production. The thin, nearly circular blanket seen in Fig. 5 covering the target around the impact area apparently results from liquid material that has covered over the impact hillocks generated earlier.

In conclusion, there are several indications of a transient fluidization of the impacted material: the planarized areas of eroded polycrystalline diamond films, hillock instead of crater formation, and the fluidized blanket around larger impact structures. The cluster impact induced nanomodification thus seems to be governed by the localized collective motion of the impacted material.

ACKNOWLEDGMENTS

The authors thank L. P. Biró for stimulating discussions and A. Maier for providing the diamond films.

¹P. R. W. Henkes and R. Klingelhöfer, *J. Phys. (France) Colloq.* **50**, 159 (1989).

²J. Gspann, *From Clusters to Crystals*, edited by P. Jena, S. N. Khanna, and B. K. Rao (Kluwer Academic, Dordrecht, 1992), Vol. 2, p. 1115.

³J. Gspann, *Nucl. Instrum. Methods Phys. Res. B* **112**, 86 (1996).

⁴C. L. Cleveland and U. Landman, *Science* **257**, 355 (1992).

⁵Z. Insepov, M. Sosnowski, and I. Yamada, *Trans. Mater. Res. Soc. Jpn.* **17**, 111 (1994).

⁶J. Gspann, *Surf. Rev. Lett.* **3**, 897 (1996).

⁷J. Gspann and H. Vollmar, *J. Chem. Phys.* **73**, 1657 (1980).

⁸J. Gspann, *Sens. Actuators A* **51**, 37 (1995).

⁹R. B. Baldwin, *Phys. Earth Planet. Interiors* **6**, 327 (1972).

Molecular patterns by manipulating DNA molecules

Zhen-qian Ouyang, Jun Hu, Sheng-fu Chen, Jie-lin Sun, and Min-qian Li^{a)}
Shanghai Institute of Nuclear Research, Chinese Academy of Sciences, P.O. Box 8204,
Shanghai 201800, China

(Received 8 September 1996; accepted 1 April 1997)

Manipulating DNA molecules to form molecular patterns on a nanometer scale is a subject with wide prospects. By applying a modified "molecular combing" technique and imaging in air with atomic force microscope, we aligned DNA molecules on a mica surface which was chemically modified with a small organic molecule, (3-aminopropyl)-triethoxysilane. Two-dimensional patterns of DNA molecules were also constructed. © 1997 American Vacuum Society. [S0734-211X(97)02804-7]

I. INTRODUCTION

Imaging biological macromolecules and supermolecules with atomic force microscopy (AFM) can provide new biological knowledge because of its high resolution for biological samples.¹⁻⁴ Additionally, manipulation of individual biological macromolecules is a field of study that is not only significant but also very interesting. The methods of extending and manipulating DNA for various purposes have been reported as follows. Large DNA molecules can be stretched out by flow forces in molten agarose and then rapidly fixed by agarose gelation without application of electrical fields.⁵ In order to study the conformational dynamics of an individual DNA molecule, a new technique termed the "pulse-oriented field" method was adopted.⁶ Bead DNA based optical tweezers were used to measure the elastic properties of individual double-stranded and single-stranded DNA molecule.^{7,8} In addition, by first anchoring one end of the molecule in a matrix, DNA molecules were extended and aligned by a receding air-water interface or by electrophoresis. To anchor DNA to a glass surface, researchers first grafted a monolayer of silane molecules onto a substrate surface,⁹ then stretched it by a "molecular combing" technique or by gel electrophoresis.¹⁰⁻¹³

Extension of DNA is much more important for imaging with AFM which is expected to have potential application in biology. For example, direct mapping of genes or direct sequencing of DNA with AFM requires that the DNA strands be straightened first.¹⁴⁻¹⁶ To directly observe the process of reaction between DNA and other biomolecules, the DNA strands should not be tangled. However, extending DNA for AFM imaging must fulfill special requirements. Usually, biological samples must be strongly attached to an atomic flat matrix surface so that they are immobile in a buffer solution and not swept away during imaging with AFM. Although the mica surface is flat at the atomic level, unfortunately, it is not strong enough to attach biosamples.^{17,18} Chemically modified mica, particularly mica silanized with a monolayer film of (3-aminopropyl) triethoxysilane (APS), has shown good performance.^{19,20} These surfaces have the following charac-

teristics: (i) they meet the requirement of being molecularly or atomically flat; (ii) they have a strong binding ability to DNA; (iii) their preparation is simple and convenient; (iv) they are stable in a buffer solution for a long time.²¹ More important, the molecular combing technique, as well as the modified method by Weier,²² has made it possible to extend large size DNA strands into an aligned form on APS mica.

In this article, we report that large size DNA molecules could not only be aligned in one direction but also constructed into a special two-dimensional pattern on the APS mica surface.

II. MATERIAL AND METHOD

The mica surfaces were chemically modified, using a monolayer film of 3-aminopropyl triethoxysilane (APS, United Chemical Co., Bristol, PA) with a molecular self-assembly method. The detailed method was described in Ref. 10.

Supercoiled lambda DNA molecules were linearized with restriction endonuclease Nae I, both incubated for 60 min at 37 °C. Lambda DNA (48 502 bp) and Nae I were products of Promega Co., USA. Nae I recognizes the special sequence (GCC↓GGC) and produces blunt ends. There is only one recognition site in the lambda DNA molecule to restrict enzyme Nae I. The DNA digested by Nae I was diluted to a concentration of a few nanograms per microliter. For the study with AFM, the samples were prepared by first depositing a small drop (typically 2 μ l) of DNA solution onto a clean glass cover slip. The glass cover slip was then carefully placed onto the top of the APS film. The weight of the thin cover forced the solution to spread immediately into a thin layer. After a few minutes, the glass slip was removed and the APS film surface was rinsed with doubly distilled water and dried with a flow of clean nitrogen.

In this study, all images were collected using a Nanoscope III AFM (Digital Instruments, Inc., Santa Barbara, CA) with tapping mode in air.

III. RESULTS AND DISCUSSIONS

Figure 1(a) showed the typical one-dimensional DNA pattern. In this AFM image, the strands were well distributed on the APS film in a parallel fashion. The condition to form a

^{a)} Author to whom correspondence should be addressed; Electronic mail: mqli@fudan.ac.cn

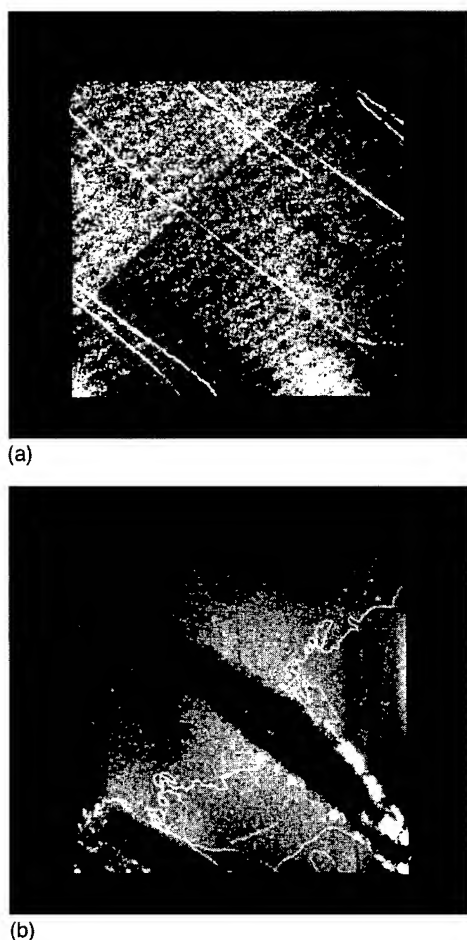


FIG. 1. AFM images of one-dimensional DNA strands aligned on the APS film surface by applying a molecular combing technique. (a) Scanning range: $10\ \mu\text{m} \times 10\ \mu\text{m}$, z range: 0–2 nm (b) DNA strands on APS film with flaws. Scanning range: $1.1\ \mu\text{m} \times 1.1\ \mu\text{m}$, the z range: 0–6 nm.

very straight DNA pattern was very subtle. Sometimes the roughness of the substrate surfaces influenced the spreading of the solution and resulted in many tangled DNA strands.²³ We found that the DNA strands could be made straighter by applying extra pressure with a finger on the glass cover slip while combing. Here, we refer to the pressure from the finger as an extra force. The reason for the pressure effect might be as follows. DNA solution can be spread by the gravitational force (the weight of the glass slip itself) and the wetting force between the two surfaces with water inside. If the spreading speed was very slow, the DNA molecules had more opportunity to be adsorbed in a small area on the APS mica surfaces, resulting in tangled fashions. The extra force could enhance the spreading speed of the solution and the straighter DNA patterns could be easily obtained.

In some cases the mica surface was not fully covered by APS film as shown in Fig. 1(b). The black flaw is the bare mica surface since the APS only forms a monolayer on mica substrates. The thickness of the APS film can be measured from the height difference between the APS area and the bare mica area. In this experiment it was about 0.6 nm. It was thought that the APS molecule attached chemically on a

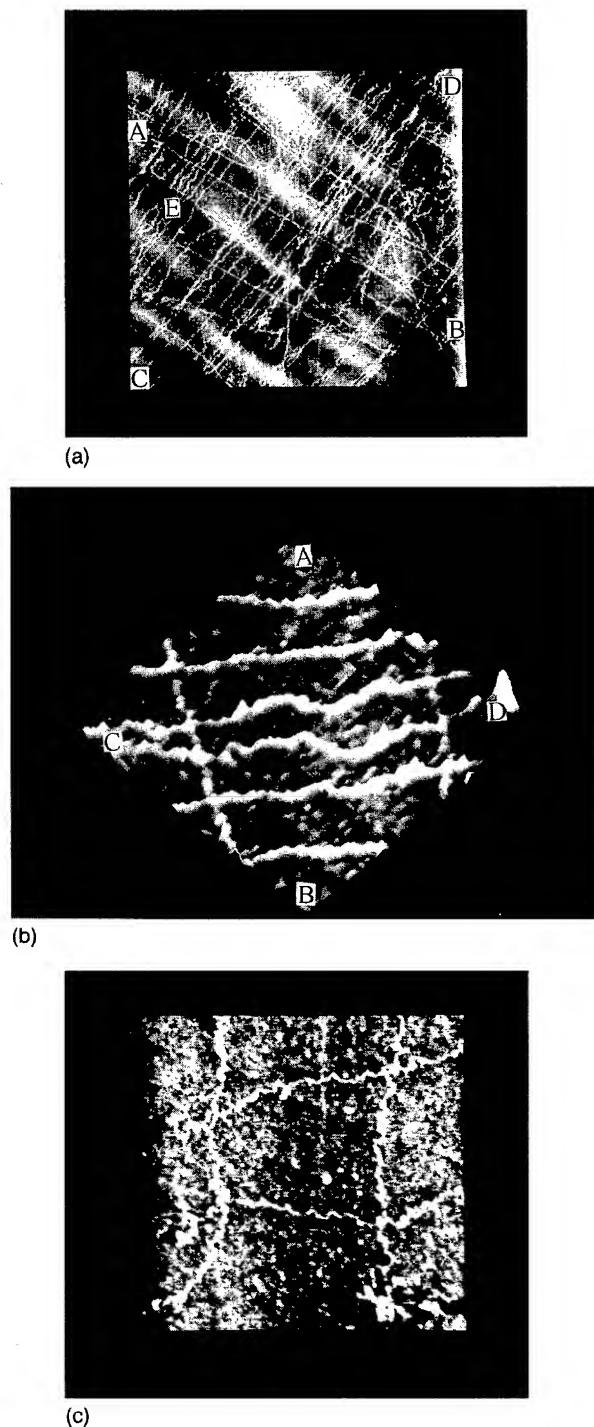


FIG. 2. AFM images of two-dimensional DNA pattern aligned on the APS film. (a) Two-dimensional DNA network in a large scanning range: $10\ \mu\text{m} \times 10\ \mu\text{m}$, (b) highlight view extracted from the E area in (a), and (c) a typical Chinese character well.

mica surface by an oxygen bridge between a silicon atom and the mica surface.^{9,19} Theoretically, the monolayer film thickness can be obtained by calculating the sum of all relative covalent bond lengths. It is about $1.0 \pm 0.1\ \text{nm}$.²⁴ The 0.6 nm thickness measured by AFM here might come from the fact that the film is soft and may easily be flattened by the

AFM tip. Another possibility is that the APS molecule is not fully upright on the mica surface.

We noticed that the DNA strands only extended over the APS film and did not enter the flaw where the mica was not covered by APS, illustrating the different adsorption of DNA on different surfaces. The adsorption between DNA and the mica surface is weaker than that between DNA and the APS film.

Figure 2 contains AFM images of two-dimensional DNA molecule patterns. To construct the patterns, we used the following procedures. First of all, a monolayer of APS film was grafted on a mica surface; second, the DNA molecules were aligned on it in the AB direction, rinsed, and dried; then we aligned the DNA along the CD direction, rinsed, and dried again. This result shows that the APS film is a significant contributor to aligning DNA and constructing the DNA network. Figure 2(a) is a two-dimensional network on a large scale. Figure 2(b) is a highlight view in a small range extracted from the *E* area in Fig. 2(a). Figure 2(c) is a typical Chinese character "well."

In view of the strong adsorption, the DNA strands aligned on the APS film were difficult to rinse away, therefore, the APS mica system provided a suitable method for direct mapping of genes and direct sequencing of DNA with AFM. It is useful not only to obtain a low "background noise" sample for AFM imaging but also to prepare two-dimensional DNA networks. The mechanism of adsorption in this system is not clear. One possible mechanism might be the interaction between DNA and APS film by electrostatic attraction. The APS molecules were grafted onto the mica surface and the up ends were amine groups-NH₃⁺ in water solution at pH=7 while DNA is usually negative charged.¹⁰ Though the strength of the attraction would be very weak at each attachment site, it would be enormous for a long strand molecule because of frequent points of attachment along its full length.

IV. SUMMARY

By applying a modified molecular combing technique, we aligned DNA molecules on APS mica surfaces and imaged in air with AFM. Two-dimensional patterns of DNA molecules could be constructed and a typical Chinese character

well was formed. Alignment of DNA strand may have potential applications in constructing a molecular electric circuit and precision gene mapping.

ACKNOWLEDGMENTS

The authors like to thank Yi Zhang and Gui-gen Zhang for much help in this experiment. This work was supported by the Shanghai Center of Life Sciences, and the Major Projects of Chinese Academy of Sciences (KJ 951-A1-603).

- ¹C. Bustamante and D. Keller, *Phys. Today* **48**, 32 (1995).
- ²M.-Q. Li, L. Xu, and A. Ikai, *J. Vac. Sci. Technol. B* **14**, 1410 (1996).
- ³M.-Q. Li, H. G. Hansma, J. Vesenska, G. Kelderman, and P. K. Hansma, *J. Biomol. Struct. Dyn.* **10**, 607 (1992).
- ⁴M.-Q. Li, H. G. Hansma, G.-F. Hong, X.-W. Yao, and P. K. Hansma, *J. Vac. Sci. Technol. B* **12**, 1456 (1994).
- ⁵D. C. Swartz and C. R. Cantor, *Cell* **37**, 67 (1984).
- ⁶D. C. Swartz and M. Koval, *Nature (London)* **338**, 520 (1989).
- ⁷S. B. Smith, L. Finzi, and C. Bustamante, *Science* **258**, 1122 (1992).
- ⁸S. B. Smith, Y.-J. Cui, and C. Bustamante, *Science* **271**, 795 (1996).
- ⁹L. Netzer and J. Sagiv, *J. Am. Ceram. Soc.* **105**, 674 (1983).
- ¹⁰J. Hu, M. Wang, H.-U. G. Weier, P. Frantz, W. Kolbe, D. F. Ogletree, and M. Salmeron, *Langmuir* **12**, 1696 (1996).
- ¹¹A. Bensimon, A. Simon, A. Chiffaudel, V. Croquette, F. Heslot, and D. Bensimon, *Science* **265**, 2096 (1994).
- ¹²W. D. Volkmuth and R. H. Austin, *Nature (London)* **358**, 600 (1992).
- ¹³H. Kabata *et al.*, *Science* **262**, 1561 (1993).
- ¹⁴H. G. Hansma, A. L. Weisenhorn, S. A. C. Gould, R. L. Sinsheimer, H. E. Gaub, G. D. Stucky, C. M. Zaremba, and P. K. Hansma, *J. Vac. Sci. Technol. B* **9**, 1282 (1991).
- ¹⁵D. D. Dunlap and C. Bustamante, *Nature (London)* **342**, 204 (1989).
- ¹⁶R. J. Driscoll, M. G. Youngquist, and J. D. Baldeschwieler, *Nature (London)* **346**, 294 (1990).
- ¹⁷H. G. Hansma, *J. Vac. Sci. Technol. B* **14**, 1390 (1996).
- ¹⁸M. Bezanilla, B. Drake, E. Nudler, M. Kashlev, P. K. Hansma, and H. G. Hansma, *Biophys. J.* **68**, 1672 (1995).
- ¹⁹Y. L. Lyubchenko, A. A. Gall, L. S. Shlyyakhtenko, R. E. Harrington, B. L. Jacobs, P. I. Oden, and S. M. Lindsay, *J. Biomol. Struct. Dyn.* **10**, 589 (1992).
- ²⁰Y. L. Lyubchenko, L. S. Shlyyakhtenko, R. E. Harrington, P. I. Oden, and S. M. Lindsay, *Proc. Natl. Acad. Sci. USA* **90**, 2137 (1990).
- ²¹M. Bezanilla, S. Manne, D. E. Laney, Y. L. Lyubchenko, and H. G. Hansma, *Langmuir* **11**, 655 (1995).
- ²²H. U. G. Weier, M. Wang, J. C. Mullikin, Y. Zhu, J.-F. Cheng, K. M. Greulich, A. Bensimon, and J. W. Gray, *Hum. Mol. Genet.* **4**, 1903 (1995).
- ²³T. Thundat, D. P. Allison, and R. J. Warmack, *Nucleic Acids Res.* **22**, 4224 (1994).
- ²⁴D. R. Lide *et al.*, *Handbook of Chemistry and Physics*, 72nd ed. (Chemical Rubber, Boca Raton, FL, 1991).

Observation of modification and recovery of local properties of polyethylene oxide

H.-Y. Nie^{a)}

Joint Research Center for Atom Technology, National Institute for Advanced Interdisciplinary Research, Higashi 1-1-4, Tsukuba, Ibaraki 305, Japan

M. Motomatsu^{b)}

Joint Research Center for Atom Technology, Angstrom Technology Partnership, Higashi 1-1-4, Tsukuba, Ibaraki 305, Japan

W. Mizutani^{c)} and H. Tokumoto^{c)}

Joint Research Center for Atom Technology, National Institute for Advanced Interdisciplinary Research, Higashi 1-1-4, Tsukuba, Ibaraki 305, Japan

(Received 8 September 1996; accepted 26 February 1997)

By applying large forces to the surface during an atomic force microscopy scanning, crystallized polyethylene oxide (PEO) films were modified and showed a new structure in which there are the formation of a nanometer-scale bundle structure and an increase of elastic stiffness. The modified part was removed by dropping benzene onto a rotating surface through a peeling process for the case of thin modified layers, while for thick layers, the local properties were recovered by absorbing benzene into the modified parts. In the course of these measurements, the thickness of the modified layers was estimated. © 1997 American Vacuum Society. [S0734-211X(97)03004-7]

I. INTRODUCTION

Recently, there has been increasing interest in characterizing¹ as well as locally modifying polymers with atomic force microscopy (AFM).²⁻⁵ It was reported that bundles were formed on a polyacetylene surface due to scanning an AFM tip on the surface.² We have demonstrated that the morphology and mechanical properties of a crystallized polyethylene oxide (PEO) film with spiral and layered structures prepared on mica surfaces can be locally modified by applying forces at 20 nN during scanning in an AFM.⁵ The modification in morphology is a formation of bundles with spacing distances of about 15 nm on the PEO surface. The modification in mechanical properties is an increase in elastic modulus as estimated from the measurement of force-distance curves or in elastic stiffness as imaged by the force modulation technique.⁴⁻⁷ This modification is believed to be due to the rearrangement of the PEO molecular chains caused by the interaction between them and the AFM tip. However, as far as we know, there has been no report on the depth and surface property variation of the modified polymer structure.

The modified structure can be removed by dropping benzene, which is a good solvent for PEO, onto the rotating sample surface, from which the depth information is obtained. There are two reasons for rotating the sample during the benzene dropping action, i.e., controlling the amount of benzene absorbed into the PEO film to prevent its dissolution

and providing a centrifugal force to remove the modified PEO structure mechanically from the sample surface. Based on the experimental observations of removed parts and complete recovery of the modified PEO structures, we clarify that the depth of the modified structure is dependent on the applied forces and the change in the mechanical properties is limited to the modified PEO structures. These results are unique in contributing to the understanding of the local modification of polymers caused by tip-surface interaction in the AFM. We will also discuss the mechanism for the observed removal and recovery of the modified PEO structure by considering possible differences in benzene absorption due to structure changes of the molecular chains in the modified PEO.

II. EXPERIMENT

The crystallized film of PEO was prepared by coating a PEO solution in benzene onto a cleaved mica surface mounted on a spin coater followed by annealing the sample in a moderate vacuum at 90 °C for 24 h (PEO's melting temperature is 66 °C).⁸ The thickness of the sample was about 200 nm. To map the mechanical properties, we modified a commercial AFM system (SPA 300, Seiko Instruments Inc.) so as to oscillate the sample height with a peak-to-peak amplitude of 1 nm at 5 kHz and to record the response of the cantilever with a lock-in amplifier. The response of the cantilever to an oscillation of the sample height corresponds to the slope of the force-distance curve available after the tip mechanically contacts with the sample surface. This slope is not only determined by the elastic modulus (Young's modulus) of the sample, but also by the spring constant of the cantilever and the applied force between the tip and sample surface.⁵ The response of the cantilever to the oscillation of the sample height can qualitatively express the elastic stiff-

^{a)}Present address: Yokohama Research Center, Mitsubishi Chemical Corporation, 1000 Kamoshida-cho, Aoba-ku, Yokohama 227, Japan; Electronic mail: hynie@rc.m-kagaku.co.jp

^{b)}Present address: Advanced Materials Application Laboratory, DuPont Kabushiki Kaisha, 2-2-1 Hayabuchi, Tsuzuki-ku, Yokohama 224, Japan.

^{c)}Permanent address: Electrotechnical Laboratory, Umezono 1-1-4, Tsukuba, Ibaraki 305, Japan.



FIG. 1. An AFM image of a bundle structure of PEO crystallized surface formed by scanning a tip at 20 nN. The imaging area is $110 \times 110 \text{ nm}^2$ and the grey scale is from 0 to 15 nm.

ness for the sample surface. Using this technique we can map the elastic stiffness distribution simultaneously with topography of the sample surface. Details of the technique were described elsewhere.⁵ A rectangular shaped silicon cantilever with a spring constant of 0.75 N/m and a resonance frequency of 88 kHz (Olympus Opt. Inc.) was used in this study. The cantilever is 0.8, 40, and 100 μm in thickness, width, and length, respectively. The tip is 2.8 μm long and the radius of the apex is about 20 nm. All measurements were performed in air at room temperature with a relative humidity of 50%–60%.

Here, we describe procedures for the modification of the crystallized PEO film and for the removal of the modified PEO structure. A large area of 8 μm square was first imaged with small forces of 5 nN under which there was no detectable degradation on the surface. Then a smaller area of 5 μm square was modified by scanning the tip with large forces (for example, 50 nN). The modification was confirmed by observing both the bundle structures and the increase in the elastic stiffness. To probe the depth of and surface property variation on the modified PEO structure, we replaced the sample onto the spin coater and dropped benzene onto the sample surface while rotating at a rate of 1000 rpm. Although the accurate quantity of benzene absorbed into the PEO films could not be estimated, we gave a rough qualitative description for the amount of benzene in terms of the number of drops of benzene. Then the sample was placed in a moderate vacuum at room temperature for 12 h to evacuate residual benzene. After the above treatments the sample was remounted on the AFM scanner for topography and elastic stiffness measurements.

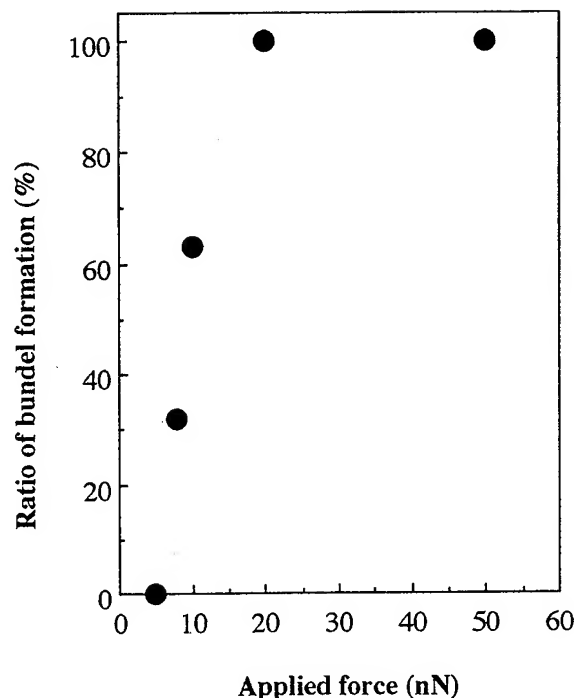


FIG. 2. The ratio (%) of the area with bundle structure to the entire area scanned by a tip at various forces.

In order to exactly observe the same area, we prepared a marker by scratching crossed lines on the backside surface of the mica substrate and placed the AFM tip around there with the help of an optical microscope through the transparent mica substrate and thin PEO film.

III. RESULTS

Figure 1 shows a typical AFM image (110 nm square) of a bundle structure on the PEO surface formed by scanning the AFM tip at 20 nN. The structure is seen as regular white bundles perpendicular to the scanning direction (indicated by an arrow) of the large force scanning. The spacing between the bundles is about 15 nm and the corrugation height is about 3 nm. The length of the bundles is on the order of 100 nm.

We found that there was a threshold force for the formation of bundles when scanning the AFM tip on the PEO surface. We imaged (at 5 nN) a series of topography of areas modified at various forces and calculated the ratio of the bundled area to the entire area that experienced the large force scanning. The result is summarized in Fig. 2, showing the existence of a threshold at ~ 10 nN. We noted that the surface of PEO showed an aging effect in the modification, that is, a fresh surface needs small applied forces while an aged one needs larger applied forces for being modified. For example, aged PEO surfaces kept in air for about one month could never be modified even at the applied forces as large as 50 nN. This effect is tentatively attributed to a lubricant effect on the PEO surface due to adsorption of water.⁹

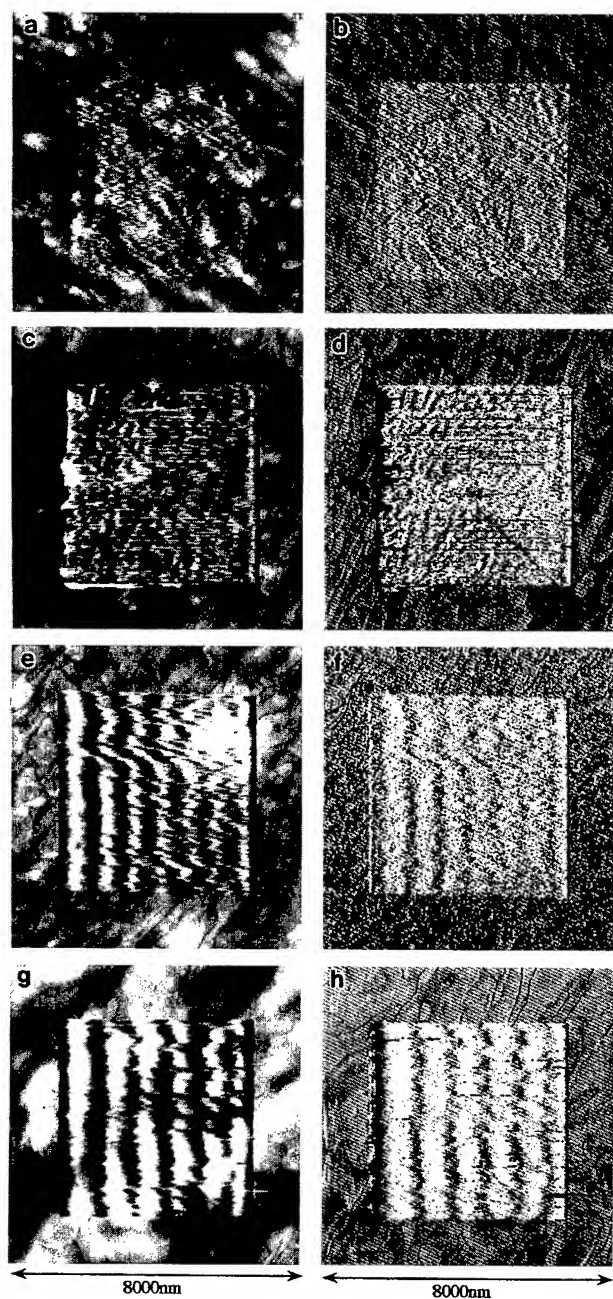


FIG. 3. Simultaneously obtained topographic (left-hand side) and elastic stiffness distribution (right-hand side) images of PEO surfaces modified by scanning a tip at large forces of 25, 50, 100, and 150 nN at the central areas ($5000 \times 5000 \text{ nm}^2$). The grey scales (from the dark to the white of the image) for the topographic images of (a), (c), (e), and (g) are from 0 to 56, 85, 95, and 174 nm, respectively. Data in elastic stiffness images are a direct record from a lock-in amplifier of the response of the cantilever to the oscillation of the sample height at 5 kHz with a peak-to-peak amplitude of 1 nm. The average increase of the elastic stiffness on the modified PEO shown in (b), (d), (f), and (h) is 1.8%, 2.8%, 3.0%, and 3.3%, respectively.

We extensively studied the modified PEO structure as a function of the applied forces above the threshold force. Figure 3 shows the results of modification in topography (left-hand side) and elastic stiffness (right-hand side) of PEO surfaces by scanning an AFM tip over 5000 nm square at 25 [(a), (b)], 50 [(c), (d)], 100 [(e), (f)], and 150 [(g), (h)] nN,

respectively. All images were obtained at 5 nN and the sample height was oscillated in a peak-to-peak value of 1 nm for imaging the elastic stiffness distribution.

The morphology of all of the modified areas in Fig. 3 is seen as being scratched by the tip. The modified surfaces in Figs. 3(a) and 3(c) are seen as a hammered-out plate, while those in Figs. 3(e) and 3(g) show an undulation character. When the applied force is increased, the undulation character becomes clear. As shown in Figs. 3(e) and 3(g), there are seven and six bundles in the area of 5000 nm square, respectively. This micron-scale bundle structure is similar to that reported in Ref. 3, whose spacing distance is dependent on the applied forces. By magnifying all of the modified areas shown in Figs. 3(a), 3(c), 3(e), and 3(g), we confirmed the existence of nanometer-scale bundles perpendicular to the scanning direction (not shown), which is similar to that shown in Fig. 1. Although the modified PEO surfaces caused by various applied forces show different surface morphology on a micron-scale, there is an identical nanometer-scale bundle formation on all of the modified areas shown in the topographic images in Fig. 3. Therefore, two types of structure exist on the modified PEO surface. One of them is a bundle structure with spacing distances of about 15 nm, which is formed when scanning a tip at forces over the threshold value of $\sim 10 \text{ nN}$. The other one is the micron-scale bundle structure which is formed when scanning tip at forces over about ten times of the threshold value. We noted that the spacing distance of the micron-scale bundles is dependent on the applied forces while that of the nanometer-scale bundles is not.

The simultaneously obtained elastic stiffness distribution is shown on the right-hand side in Fig. 3 in which the elastic stiffness is expressed in a manner where the brighter contrast corresponds to the larger elastic stiffness. It is clear that there is an increase in the elastic stiffness on the modified areas as shown in Figs. 3(b), 3(d), 3(f), and 3(h). The elastic stiffness imaging is thus effective to qualitatively measure the elastic stiffness distribution on a surface. For a quantitative analysis on the elastic modulus, however, force-distance curves should be measured.⁵ In this article we only show a change in the elastic stiffness due to the modification. From cross sectional analysis for the images shown in those figures, we found a relative increase of 1.8%, 2.8%, 3.0%, and 3.3% in the elastic stiffness on the modified areas caused by an applied force of 25, 50, 100, and 150 nN, respectively.

Figures 4(a) and 4(c) show topographic images for the samples modified at 25 and 50 nN and then treated by dropping five and ten drops of benzene onto the surface as described in the experimental section, respectively. The images shown in Figs. 4(a) and 4(c) are exactly the same areas as those for Figs. 3(a) and 3(c), respectively. It is clear that certain parts of the modified PEO structures are removed and the underlying PEO surfaces are visible. We also see that some single bundles still remained in Fig. 4(a) but not in Fig. 4(c). There are no bundles detected on the newly appeared underlying PEO surfaces that look to be identical to the pristine PEO surfaces. The simultaneously obtained elastic stiff-

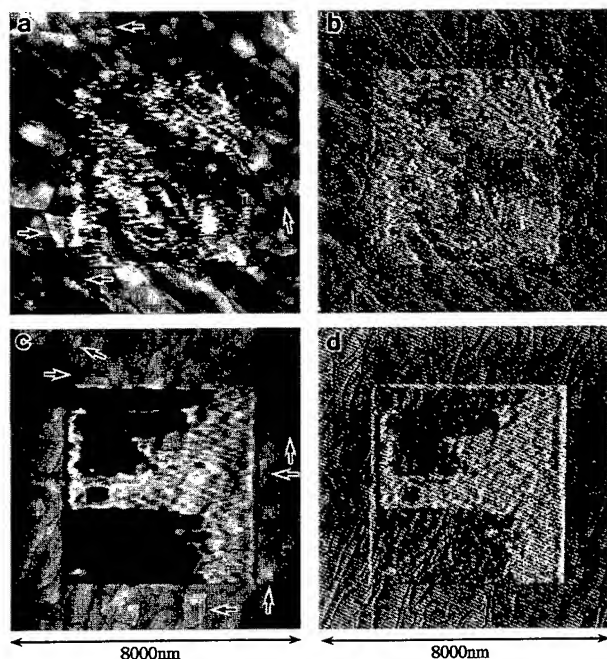


FIG. 4. Simultaneously obtained topographic [(a) and (c)] and elastic stiffness distribution [(b) and (d)] images of the 25 and 50 nN modified PEO surfaces at the same locations as those in Figs. 3(a) and 3(c), respectively, after benzene dropping action. Arrows in (a) and (c) indicate slight reduction of edges by benzene. In (b) and (d) the newly appeared underlying PEO surfaces exhibit the same elastic stiffness as the pristine ones and some of the modified structures still remain and show a larger elastic stiffness similar to that in Figs. 3(b) and 3(d), respectively. The grey scale for the topographic images in (a) and (c) are from 0 to 56 and 125 nm, respectively.

ness distribution images in Figs. 4(b) and 4(d) show that the remaining modified PEO parts have larger elastic stiffness than the underlying PEO. It also shows that the elastic stiffness distribution on newly appearing underlying PEO surfaces is the same as that on the pristine PEO, indicating that the elastic stiffness change is limited to the modified layers.

This type of experiment for removing the modified PEO parts is useful to obtain depth information for the modification. From cross sectional analysis of the images shown in Figs. 4(a) and 4(c), we estimated average height differences of about 10 and 40 nm, respectively, between the remaining modified parts and the newly appeared underlying PEO surfaces. From these results, we find that the thickness of the modified layer at 25 nN is about 10 nm, which corresponds to the thickness of the topmost layer of the crystallized PEO. When the applied force increased to 50 nN, the PEO film was modified into a depth of about 40 nm, corresponding to four layers of PEO. The removal of the modified PEO structures caused by relatively smaller applied forces, especially by 25 nN, shows a good reproducibility. Based on the morphology as shown in Figs. 4(a) and 4(c) the removal process of the modified PEO is considered as a mechanical peeling of the modified PEO layers during dropping benzene onto the rotating surface.

By carefully comparing Figs. 3(a) and 3(c) with Figs. 4(a) and 4(c), we note that there are some small changes in the morphology of the pristine PEO after the benzene dropping

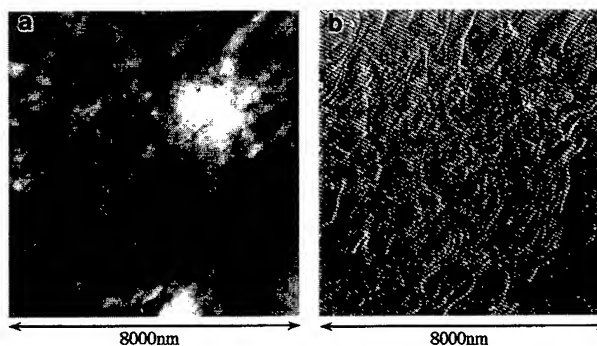


FIG. 5. Simultaneously obtained topographic (a) and elastic stiffness distribution (b) images of the 100 nN modified PEO surface at the same location as that in Fig. 3(e) after dropping 20 drops of benzene onto the rotating surface. The modified PEO surface was recovered. The grey scale for the topographic image is from 0 to 95 nm. No elastic stiffness difference between the pristine and the once-modified surface was detectable.

action, i.e., the edges of layered structures are observed to be removed slightly as indicated by arrows in Figs. 4(a) and 4(c). As the amount of benzene dropped onto the surface increases from five to ten drops, by comparing Figs. 4(a) with 4(c) we can see that the reduction of edges becomes clear. The results shown in Figs. 4(a) and 4(c) indicate that the removal of the pristine PEO caused by the benzene dropping action is limited to the edges. This experimental result indicates that the rotating surface of the crystallized PEO film cannot absorb benzene effectively, while the edges of the layered structures can trap some benzene. Therefore, when we drop benzene onto the rotating PEO sample surface, benzene starts to attack the pristine PEO from the edges but not surfaces.

For the modified PEO structures under applied forces of 100 and 150 nN shown in Figs. 3(e) and 3(g), respectively, we did not observe any removal of the modified PEO structures by dropping ten drops of benzene onto the rotating surface. When we increased the number of benzene drops, we observed a recovery of the modified PEO structures to layered crystalline structures. Figures 5(a) and 5(b) show resulting changes in both topography and elastic stiffness distribution after dropping 20 drops of benzene at the same area shown in Fig. 3(e). In the topographic image we can see some footprint, whereas in the elastic stiffness image, no contrast difference between the pristine and the once-modified PEO is detectable. It is clear that the modified area is recovered to being similar to the pristine one. This recovery process is considered to be due to a swelling effect of molecular chains in the modified PEO, which is a relaxation process of polymer molecular chains upon absorbing solvents. We noted that the benzene dropping action has a relatively weak influence on the pristine part of PEO, which may have originated from the difference in surface structure for the modified and pristine PEO. When benzene is absorbed adequately into the modified PEO layers, the molecular chains thoroughly swell, resulting in a recovery both in surface morphology and elastic stiffness. The experimental result strongly indicates that the modified parts of PEO directly

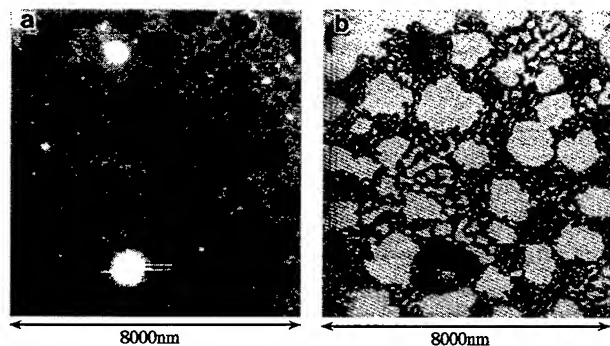


FIG. 6. Simultaneously obtained topographic (a) and elastic stiffness distribution (b) images after dropping 40 drops of benzene onto the rotating surface. The PEO film was dissolved by benzene and only a monolayer PEO network remained. The mica substrate surface is visible and shows an elastic stiffness of 10% larger than that of the PEO. The grey scale for the topographic image is from 0 to 80 nm. Data in elastic stiffness image are a direct record from a lock-in amplifier of the response of the cantilever to the oscillation of the sample height at 5 kHz with a peak-to-peak amplitude of 1 nm.

contribute to the increase of the elastic stiffness, which may be caused by the compact molecular chains arrangement due to the pressure induced by the tip.

When we dropped 40 drops of benzene onto the surface, we observed that the PEO films were finally dissolved into benzene and the surface of the mica substrate appeared as shown in Fig. 6. In Fig. 6(a) only a monolayer network of PEO molecules with a thickness of about 10 nm remains and the remaining parts are assigned to the mica substrate due to the larger elastic stiffness⁴ expressed in Fig. 6(b) as a contrast of 10% larger than that on the PEO films. We confirmed that dropping several drops (for example, five drops) onto a nonrotating surface also caused a result similar to that in Fig. 6, which indicates that a rotating surface hinders the absorption of benzene into the film and hence results in the mechanical peeling.

We also confirmed that after annealing the sample at 90 °C (the temperature at which we prepared the sample), the modified area was completely recovered and there was no detectable difference in both morphology and elastic stiffness between the pristine PEO and the once-modified surfaces. At a high temperature of 90 °C, PEO melts and recrystallizes so that the modified and pristine molecular chains could not be distinguished.

IV. DISCUSSION

The relationship between the depth of the modified PEO layers and the applied forces is revealed through the peeling process of the modified PEO parts obtained by dropping benzene to the rotating surface. The PEO layers near the surface can be modified by the tip pressure. When the applied force is 25 nN, only the topmost layer is modified. When the force increases to 50 nN, four layers can be modified. This means that when the applied force increases to 50 nN, it can influence the PEO films up to four layers. Since the contact area of the tip and the sample surface is small, the pressure can be very large, e.g., 0.1 GPa if we assume a contact area of

500 nm², which is calculated from the Hertzian contact between the tip with a radius of 20 nm and PEO film with a Young's modulus of 0.3 GPa⁵ under an applied force of 50 nN. Under a large pressure, the PEO molecular chains can be rearranged to form bundles and hence to contribute at least partly to the increase of the elastic stiffness. For the modification by forces of 100 and 150 nN as shown in Figs. 3(e) and 3(g), more than four layers of the PEO layers are expected to be modified. However, in this case we could not peel off the modified region so the exact depth of the modified PEO could not be obtained. Here we discuss a possible model for the behavior in the removal of the modified PEO layers as a function of applied forces.

In our experiment we changed the amount of benzene dropped onto the surface while keeping the sample rotation speed constant at 1000 rpm. This means that we controlled the interaction time between benzene and the surface. When ~20 drops of benzene were dropped, molecular chains in the modified PEO are easier to swell than in the pristine PEO as seen in Fig. 5, indicating a difference in absorbing benzene for the two kinds of PEO structures. This difference is thus considered to result in different interaction times for the modified and pristine layers. The change in surface structure caused by tip-surface interaction during the large force scanning should correspond to the different interaction time. For the modified PEO surface, the modification possibly induces cracks or defects, which may serve as effective paths for benzene to penetrate into the PEO films so that benzene has sufficient time to swell molecular chains. For the pristine PEO surface which is crystallized, there are almost no paths for benzene to penetrate into the PEO films on the time scale under consideration. Thus benzene dropped could not have enough time to be effectively absorbed from the surface but could be trapped by the edges of the layered or spiral structures, resulting in slight removal of edges to be observed as indicated by arrows in Figs. 4(a) and 4(c).

By dropping benzene onto the rotating sample surface, we observed a peeling of certain parts of the modified PEO layers for those caused by 25 and 50 nN. Using the concept of the interaction time of benzene and the surface as discussed above, we explain the variation in peeling of the modified PEO as a function of the applied forces. For the modified PEO layers caused by 25 and 50 nN, the interaction time corresponding to dropping five or ten drops of benzene allows benzene to penetrate to the interface of the modified and underlying PEO layers to weaken the interaction there but not to swell all of the modified PEO. Because the sample was rotating at 1000 rpm during the benzene dropping action, the weakened parts of the modified PEO layers were peeled off by the centrifugal force due to the rotation.

However, we did not observe any peeling processes for the modified layers made by 100 and 150 nN in our experiment. For these modified PEO layers, because of a deeper modification, it needs a much larger amount of benzene to penetrate to the interface of the modified and underlying layers to swell the molecular chains or weaken the interaction between the modified and the underlying PEO layers. The

interaction time corresponding to dropping five or ten drops of benzene may be too short to allow benzene to penetrate to the interface. When increasing the interaction time to that corresponding to dropping 20 drops of benzene, the modified layers near the surface may start to swell and recover by absorbing benzene, which in turn prevents benzene from further penetrating to the interface. Therefore, no peeling of the deeply modified PEO layers can be realized. It is obvious that when the interaction time is extremely large the modified PEO and also the pristine PEO will be dissolved into benzene, as shown in Fig. 6. The condition for that is to drop enough of an amount of benzene (for example, 40 drops) onto the rotating surface. An alternative way to make the PEO films dissolve into benzene is to drop a small amount of benzene (for example, five drops) onto the nonrotating surface.

V. CONCLUSIONS

By applying five and ten drops of benzene onto rotating sample surfaces at 1000 rpm, we removed parts of the PEO layers modified by applied forces of 25 and 50 nN and estimated depths as 10 and 40 nm, respectively. In contrast, under the same conditions of benzene dropping action, we could not remove the PEO layers modified by forces of 100 and 150 nN. When dropping 20 drops of benzene, the modified parts completely recovered by absorbing benzene into the modified PEO parts. By simultaneously imaging elastic stiffness distribution with topography, we confirmed that the elastic stiffness change due to the modification is limited to the modified structure. An adequate amount of benzene dropped into the PEO finally results in complete dissolution of both the modified and pristine PEO surfaces so that the mica substrate surface appeared.

A model is proposed to explain the observed removal and recovery of the modified PEO structures by considering whether benzene can penetrate to the interface between the modified and underlying layers. For relatively thin modified layers ranging from 10 to 40 nm, benzene can penetrate to the interface of the modified and underlying PEO layers to weaken the interaction there. Then the weakened modified PEO parts can be peeled off by the centrifugal force due to the rotation. For the modified layers with a thickness larger

than 40 nm, benzene can swell the layers near surface before it penetrates into the interface and recovers the modified PEO parts. This process in turn prevents benzene from penetrating further into the underlying modified PEO layers. Therefore, no peeling process can be realized, at least under the condition in our experiment.

The results reported are expected to contribute to a better understanding of the modified PEO layers by investigating not only the changes in the structure and mechanical properties, but also the thickness as a function of the applied forces. However, much work remains to be done to completely understand the local change in the modified structure. For example, in the PEO layers modified by relatively large forces of 150 nN [Fig. 3(g)] the valley is softer than the protruding bundle, which indicates that the molecular chains become compact in the protruding bundle but not in the valley. Furthermore, a quantitative analysis of changes in the elastic modulus in the modified PEO surface as a function of applied forces is highly required.

ACKNOWLEDGMENTS

The authors are grateful to Professor C. F. Quate for his stimulating discussion on the modified PEO surfaces by an AFM tip. One of the authors (H.Y.N.) thanks Dr. J. Masai for encouragement and acknowledges a fellowship supported by the Research Development Corporation of Japan (JRDC). This work was supported in part by the New Energy and Industrial Technology Development Organization (NEDO).

¹S. N. Magonov, *Polym. Sci. Ser. B* **38**, 34 (1996).

²O. M. Leung and M. C. Goh, *Science* **255**, 64 (1992).

³Z. Elkaakour, J. P. Aimé, T. Bouhacina, C. Odin, and T. Masuda, *Phys. Rev. Lett.* **73**, 3231 (1994).

⁴H.-Y. Nie, M. Motomatsu, W. Mizutani, and H. Tokumoto, *J. Vac. Sci. Technol. B* **13**, 1163 (1995).

⁵H.-Y. Nie, M. Motomatsu, W. Mizutani, and H. Tokumoto, *Thin Solid Films* **273**, 143 (1996).

⁶P. Maivald, H.-J. Butt, S. A. C. Gould, C. B. Prater, B. Drake, J. A. Gurley, V. B. Elings, and P. K. Hansma, *Nanotechnology* **2**, 130 (1991).

⁷M. Radmacher, R. W. Tillmann, M. Fritz, and H. E. Gaub, *Science* **257**, 1900 (1992).

⁸M. Motomatsu, H.-Y. Nie, W. Mizutani, and H. Tokumoto, *Polymer* **37**, 183 (1996).

⁹B. H. Cao and M. W. Kim, *Europhys. Lett.* **29**, 555 (1995).

Scanning probe nanofabrication of chemically active areas on substrate covered with organosilane monolayers

Hiroyuki Sugimura^{a)} and Nobuyuki Nakagiri^{b)}

Tsukuba Research Laboratory, Nikon Co., 5-9-1 Tokodai, Tsukuba 300-26, Japan

(Received 12 September 1996; accepted 30 January 1997)

An organosilane monolayer self-assembled onto Si substrates was patterned using a scanning tunneling microscope or an atomic force microscope with a conductive probe. The tip-scanned area was etched by wet etching, while the unscanned area remained resistant to the etching. Consequently, the tip-scanned patterns were transferred to the substrate Si using the scanned area as an etching window. Moreover, the tip-scanned area preferentially reacted with other organosilane molecules, e.g., 3-aminopropyltrimethoxysilane, to form a monolayer confined to the scanned pattern. Next, onto this amino-terminated monolayer, aldehyde-modified latex nanoparticles were immobilized area selectively through chemical reaction between the amino and aldehyde groups. It was demonstrated that in the areas where the probe tip of the microscopes had been scanned, the surface was terminated with hydroxyl groups and became active to various chemical treatments.

© 1997 American Vacuum Society. [S0734-211X(97)10904-0]

I. INTRODUCTION

Self-assembled monolayers (SAMs), which spontaneously form on certain substrates in solution or vapor of precursor organosilane molecules, are patternable by scanning probe microscopes (SPMs), i.e., scanning tunneling microscope (STM), or atomic force microscope (AFM).¹⁻⁵ Such monolayers have also attracted much attention as resist materials for photon, electron, or ion beam lithography.⁶⁻¹¹ Patterned SAMs have been used as templates for the fabrication of micro- and nanostructures made up of a variety of materials through various chemical modifications, including mineralization,¹¹ immobilization of protein, fluorescent materials,^{3,6,8,12} or biological cells,^{13,14} electroless plating,^{1,6,7,15} monolayer deposition,^{6,16} and chemical vapor deposition.^{17,18}

For the application of SPM-patterned SAMs as templates for the chemical synthesis of nanostructures, it is crucial to clarify the patterning mechanism and to control the chemical properties of the patterned surface. We have confirmed that, when an alkylsilane monolayer on a Si substrate was scanned in air or another humidified atmosphere by a SPM tip, the monolayer was removed in the tip-scanned area due to tip-induced electrochemistry between the monolayer and the adsorbed water and/or oxygen.¹⁹

In this article, we report on the chemical properties and the reactivities toward chemical reagents of this tip-scanned area. The area was hydrophilic due to hydroxylation and was selectively etched in an aqueous solution of $\text{NH}_4\text{F}/\text{H}_2\text{O}_2$ more efficiently than the unscanned area. In addition, the tip-scanned area was chemically reactive with other organosilane molecules, while the unscanned area remained unreactive. Consequently, a SAM of 3-aminopropyltrimethoxysilane (APS), confined to the scanned pattern, was fabricated.

II. EXPERIMENT

Substrates [p -Si (100) wafers with a resistivity of 0.1–0.2 $\Omega\text{ cm}$] were cleaned by one of two methods. The first was immersion in a 1:1:4 (weight ratio) solution of $\text{HCl}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$ at 80 °C for 10 min. The second was exposure to ultraviolet light using an excimer lamp of $\lambda=172\text{ nm}$ (Ushio Electric, UER20-172 V) in air under a reduced pressure atmosphere (10–100 Torr). Both methods were effective in cleaning the substrate and also in making the surface hydrophilic, since the surfaces became covered with a thin oxide layer terminated with hydroxyl (OH) groups. A trimethylsilyl [TMS; $(\text{CH}_3)_3\text{Si}-$] organosilane monolayer about 0.5–1.0 nm thick was prepared on the substrates by chemical vapor deposition using hexamethyldisilazane [$(\text{CH}_3)_3\text{Si}-\text{NH}-\text{Si}(\text{CH}_3)_3$] as the source material.⁴ This treatment resulted in the sample surface becoming hydrophobic.

In order to form patterns, each sample was scanned by either the probe tip of a STM or that of an AFM. The experiments were conducted using a STM (Seiko Instruments, SPI-3600) with a conductive diamond tip and an AFM (Park Scientific Instruments, Autoprobe-CP) with a conductive probe (a custom made silicon nitride probe coated with a 20 nm-thick NiCr layer).

After the tip scanning, the chemical properties and/or reactivities of the scanned areas were investigated by applying one of three different types of treatment to the samples. The first treatment was etching in a (10 g:3 g: 100 g) solution of $\text{NH}_4\text{F}:\text{H}_2\text{O}_2:\text{H}_2\text{O}$. Both Si and its oxide dissolved in the etching solution while the organosilane monolayer was relatively resistant to etching. These etched samples were observed by the same AFM in contact mode as used for the patterning using a commercially available probe (Park Scientific Instruments, Microlever).

The second type of treatment involved washing the patterned samples in a 30 wt % H_2O_2 aqueous solution for 5 min and then blowing them dry. The samples were cooled in a refrigerator, and then removed in order to be exposed to

^{a)}Present address: Department of Materials Processing Engineering, Nagoya University, Chikusa, Nagoya 464-01, Japan; Electronic mail: sugimura@otakai.numse.nagoya-u.ac.jp

^{b)}Electronic mail: nakagiri@tsukubagw.niko.co.jp

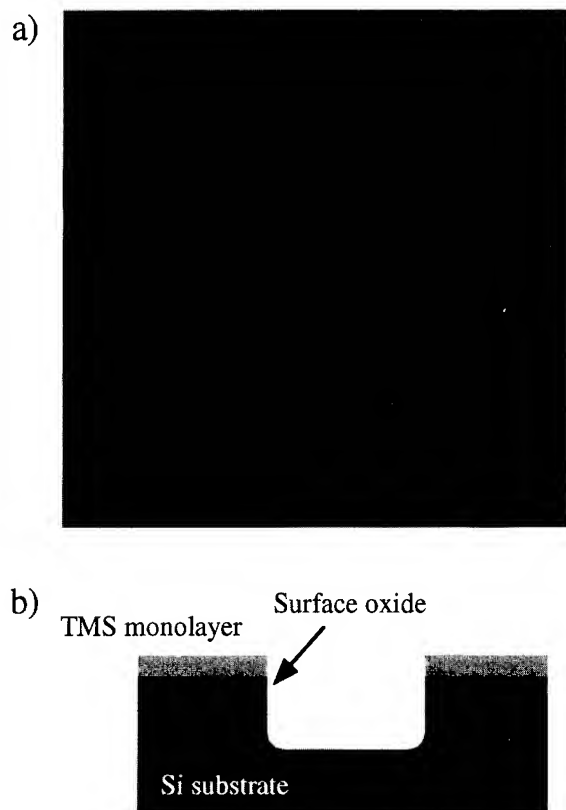


FIG. 1. Etched grooves on a Si substrate. (a) AFM image, (b) schematic cross section. A conductive AFM probe was scanned at $V_S = +10.0$ V and $v = 160$ $\mu\text{m/s}$ in air with 40% RH. Etching was conducted for 3 min. The size of the image is 10×10 μm^2 .

ambient atmosphere. Condensed water was observed by an optical microscope (Nikon, Optiphot-2).

In the third treatment, the samples were washed in distilled water and ethanol, in that order, and dried by N_2 blow. Next, they were immersed in a fresh solution of 1 vol % APS, 4 vol % CH_3COOH , and 4 vol % H_2O in methanol for 5 min. After this, the samples were rinsed in methanol twice and then distilled water. Finally, some of the APS-treated samples were immersed for 30 min in 50 ml of morpholinoethane sulfonic acid buffer (50 mM, adjusted to pH 6.5 by adding $\alpha\text{q-NH}_4\text{OH}$ solution), and 0.2 ml of an aqueous suspension of aldehyde-modified latex particles (Molecular Probes, L-5401, 2% solids in H_2O , diameter = $29 \pm 20.1\%$). These samples were again rinsed in distilled water and then blown dry. This group of samples was observed by the same AFM intermittent contact mode using a commercial probe (Park Scientific Instruments, Ultralever).

III. RESULTS AND DISCUSSION

Figures 1(a) and 1(b) show an AFM image of an etched sample and its schematic cross section, respectively. This image demonstrates that the tip-scanned area where the TMS monolayer had been degraded was selectively etched. On the contrary, the unscanned regions remained unetched due to protection by the monolayer. Grooves of 80 nm in width and 15 nm in depth were fabricated. The degree of degradation of

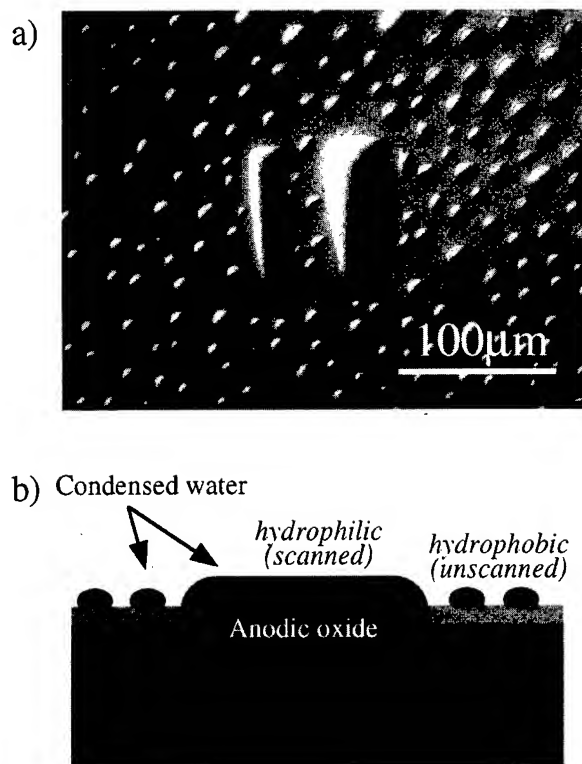


FIG. 2. (a) Optical micrograph of atmospheric water vapor condensed onto the patterned Si-TMS surface. (b) A schematic cross section. A STM tip had been scanned at $V_S = +5.0$ V in air with 40% RH over the two rectangular regions of 20×100 μm^2 and 40×100 μm^2 .

the monolayer was governed by several factors concerning the probe scanning.^{19,20} When atmospheric humidity increases, degradation proceeds more efficiently. It has been proven that oxygen also accelerates the degradation rate. When probe scan rate, v , decreases while the sample bias, V_S , remains at the same voltage, the etched grooves become deeper and wider. The degree of monolayer degradation increases with a decrease of the probe scan rate, v , that is, with an increase in the injected current per unit area. Bias voltage is also an important factor in the patterning. An increase of V_S results in an increase of junction current and, consequently, more efficient degradation of the TMS monolayer. At present, by controlling those probe scanning factors, we have achieved a maximum patterning speed of 500 $\mu\text{m/s}$ and a maximum spatial resolution of 30 nm.

Figure 2(a) is an optical micrograph showing an example of how atmospheric water vapor condensed on a tip-scanned sample cooled in a refrigerator. This is illustrated schematically in Fig. 2(b). The optical micrograph clearly indicates that the region where the tip was scanned has become hydrophilic, while the surrounding area has remained hydrophobic. The H_2O_2 rinse was not absolutely necessary for the water condensation, however, it seemed to improve the hydrophilicity of the degraded areas. It is noteworthy that, without the water condensation, we could hardly locate the scanned regions by optical microscopy. We also confirmed using the AFM that the tip-scanned area slightly protruded

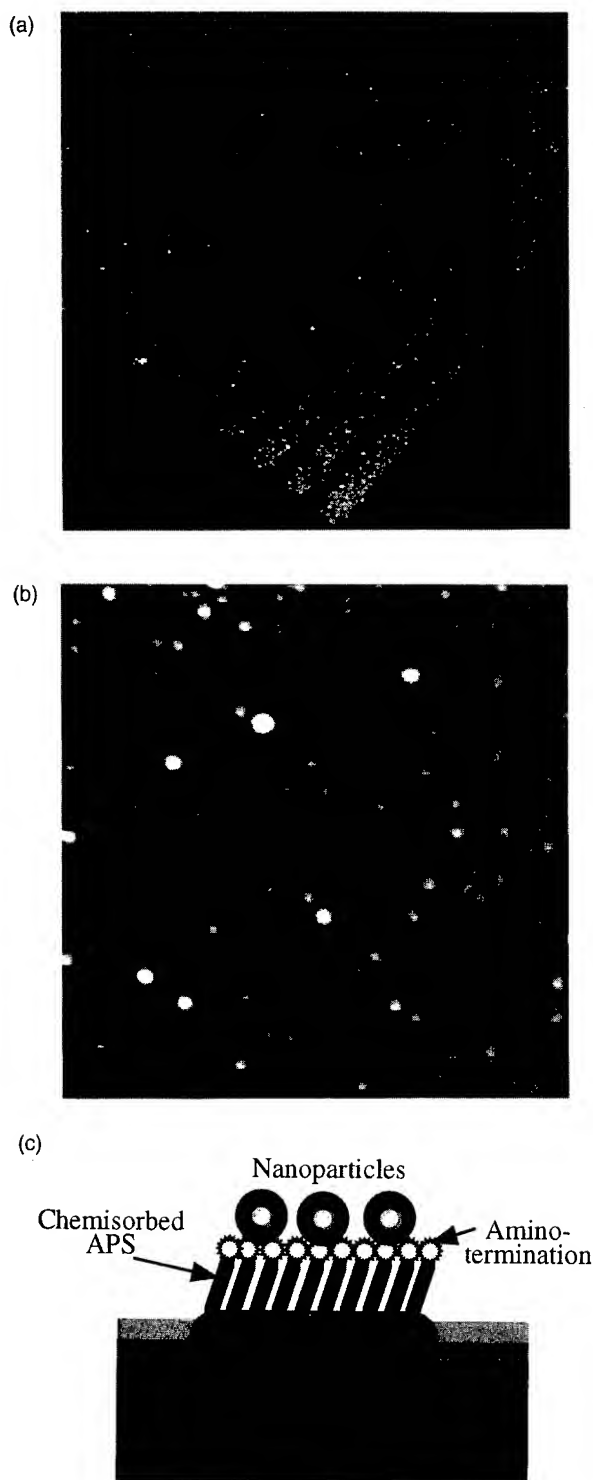


FIG. 3. (a), (b) AFM image of latex particles selectively deposited onto the tip-scanned area. (c) A schematic cross section. A STM tip had been scanned at $V_s = +5.0$ V and $v = 0.2$ $\mu\text{m/s}$ in N_2 with 80% RH. The size of the image (a) is 10×10 μm^2 , (b) is 3×3 μm^2 .

(~ 0.5 nm) from the surrounding undegraded region. This protrusion was due to the volume expansion accompanying the anodization of the substrate Si, which followed the monolayer degradation.

The AFM images in Figs. 3(a) and 3(b) show the

aldehyde-modified latex particles deposited on the tip-scanned area in a sample treated in the APS solution. There were no particles on the surrounding unscanned area. Since the thickness of the deposited-particle layer is nearly equal to the average diameter of the particles, this particle layer is thought to consist of a monolayer of the particles. It is known that the aldehyde group reacts with an amino group to form a chemical bonding as shown by Eq. (1)²¹



The particles were immobilized by this reaction between the amino groups in the APS molecule and the aldehyde groups on each particle. Therefore, APS molecules chemisorbed area-selectively onto the tip-scanned area, resulting in its termination with amino groups. In addition, we confirmed that, after the APS treatment, the tip-scanned area could not be etched in the $\text{NH}_4\text{F}/\text{H}_2\text{O}_2/\text{H}_2\text{O}$ solution. A protective coating of APS molecule had formed on the tip-scanned area.

Finally, we discuss the chemistry of the surface processed by SPM. Lercel *et al.* have reported on electron beam lithography using octadecyltrichlorosilane [OTS, $\text{CH}_3(\text{CH}_2)_{17}\text{SiCl}_3$] as resist material.^{9,10} They formed an OTS monolayer on a Si substrate and patterned it through degradation induced by electron irradiation. They reported that residue consisting of cross-linked carbon remained on the irradiated area. Marrian and colleagues have also fabricated nanoscale patterns on organosilane SAMs on Si using STM in vacuum.^{1,2} They identified their degradation mechanism as irradiation by electrons emitted from the STM tip. Except for the relatively low acceleration voltages of the electrons used in their STM patterning, this mechanism is the same as that of SAM patterning using an electron beam. However, this type of degradation mechanism does not account for our experimental results.

While there were no adsorbates on sample surfaces in vacuum, water adsorption on the surfaces of both the tip and the sample was inevitable under our experimental conditions. Therefore, it is unlikely that electrons emitted from the SPM tip propagated through the water layer and directly degraded the organosilane monolayer. The SPM tip-sample junction is connected through an adsorbed water column when the microscope is operated in the presence of atmospheric water vapor. Using such adsorbed water, we have achieved the anodization of Ti and Si surfaces.²²⁻²⁴ In a manner similar to such scanning probe anodization of metals and semiconductors, the degradation mechanism of a TMS monolayer was also concluded to be electrochemistry in the water column formed at the SPM-tip/sample junction on the basis of experimental results on humidity effect.¹⁹

We conclude that the TMS monolayer had been replaced with OH groups as a result of the tip scanning under humid conditions. This is evidenced by the selective APS chemisorption shown in Fig. 3. It is necessary that a surface onto which APS molecules are chemisorbed is terminated with OH groups, since chemisorption of organosilane molecules onto a Si oxide surface is based on the formation of siloxane bonding ($\text{Si}-\text{O}-\text{Si}$) between OH groups on the substrate sur-

face and the molecules.²⁵ Such OH termination is the origin of the hydrophilic nature of the tip-scanned area as demonstrated in Fig. 2. The amount of carbon residue on the tip-scanned surface was probably small. The alkyl chains of TMS are very short compared with those of OTS, so the removed monolayer in the case of TMS is relatively volatile.

IV. CONCLUSION

We have demonstrated nanoscale patterning of an organosilane monolayer using a STM or an AFM with a conductive probe. The TMS monolayer was selectively degraded due to tip-induced electrochemistry with adsorbed water. The monolayer was removed and the surface was terminated with OH groups. The resulting surface differs in chemical properties from the surrounding TMS monolayer and, therefore, can serve as a template for the chemical synthesis of nanostructures. The pattern of the OH-terminated surface can be transferred to a variety of nanostructures through area-selective chemical modification based on the surface differences. This is an important advantage of our system based on electrochemistry in adsorbed water and the use of a TMS monolayer.

¹C. R. K. Marrian, F. K. Perkins, S. L. Brandow, T. S. Koloski, E. A. Dobisz, and J. M. Calvert, *Appl. Phys. Lett.* **64**, 390 (1994).

²F. K. Perkins, E. A. Dobisz, S. L. Brandow, T. S. Koloski, J. M. Calvert, K. W. Rhee, J. E. Kosakowski, and C. R. K. Marrian, *J. Vac. Sci. Technol. B* **12**, 3725 (1994).

³W. T. Müller, D. L. Klein, T. Lee, J. Clarke, P. L. McEuen, and P. G. Schultz, *Science* **268**, 272 (1995).

⁴H. Sugimura and N. Nakagiri, *Langmuir* **11**, 3623 (1995).

⁵H. Sugimura and N. Nakagiri, *J. Vac. Sci. Technol. A* **14**, 1223 (1996).

⁶C. S. Dulcey, J. H. Georger, Jr., V. Krauthamer, D. A. Stenger, T. L. Fare, and J. M. Calvert, *Science* **552**, 551 (1991).

⁷J. M. Calvert, M. -S. Chen, C. S. Dulcey, J. H. Georger, M. C. Peckerar, J. M. Schnur, and P. E. Schoen, *J. Electrochem. Soc.* **139**, 1677 (1992).

⁸N. Ichinose, H. Sugimura, T. Uchida, N. Shimo, and H. Masuhara, *Chem. Lett.* 1961 (1993).

⁹M. J. Lercel, R. C. Tiberio, P. F. Chapman, H. G. Craighead, C. W. Sheen, A. N. Parikh, and D. L. Allara, *J. Vac. Sci. Technol. B* **11**, 2823 (1993).

¹⁰M. J. Lercel, M. Rooks, R. C. Tiberio, H. G. Craighead, C. W. Sheen, A. N. Parikh, and D. L. Allara, *J. Vac. Sci. Technol. B* **13**, 1139 (1993).

¹¹P. C. Rieke, B. J. Tarasevich, L. L. Wood, M. H. Engelhard, D. R. Baer, G. E. Fryxell, C. M. John, D. A. Laken, and M. C. Jachnig, *Langmuir* **10**, 619 (1994).

¹²S. Britland, E. Perez-Araud, P. Clark, B. McGinn, P. Connolly, and G. Moores, *Biotechnol. Prog.* **8**, 155 (1992).

¹³D. Kleinfeld, K. H. Kahler, and P. E. Hockberger, *J. Neurosci.* **8**, 4098 (1988).

¹⁴D. A. Stenger, J. H. Georger, C. S. Dulcey, J. J. Hickman, A. S. Rudolph, T. B. Nielsen, S. M. McCort, and J. M. Calvert, *J. Am. Chem. Soc.* **114**, 8435 (1992).

¹⁵H. Sugimura and N. Nakagiri, *Thin Solid Films* **281-282**, 572 (1996).

¹⁶N. Mino, S. Ozaki, K. Ogawa, and M. Hatada, *Thin Solid Films* **243**, 374 (1994).

¹⁷S. J. Potochnik, P. E. Pehrsson, D. S. Y. Hsu, and J. M. Calvert, *Langmuir* **11**, 1841 (1995).

¹⁸N. L. Jeon, R. G. Nuzzo, X. Xia, M. Mrksich, and G. M. Whitesides, *Langmuir* **11**, 3024 (1995).

¹⁹H. Sugimura, K. Okiguchi, and N. Nakagiri, *Jpn. J. Appl. Phys.* **35**, 3749 (1996).

²⁰H. Sugimura, K. Okiguchi, N. Nakagiri, and M. Miyashita, *J. Vac. Sci. Technol. B* **14**, 4140 (1996).

²¹A. Rembaum, *J. Immunol. Meth.* **24**, 239 (1978).

²²H. Sugimura, T. Uchida, N. Kitamura, and H. Masuhara, *Jpn. J. Appl. Phys.* **32**, L553 (1993).

²³H. Sugimura, N. Kitamura, and H. Masuhara, *Jpn. J. Appl. Phys.* **33**, L143 (1994).

²⁴H. Sugimura, T. Uchida, N. Kitamura, and H. Masuhara, *J. Phys. Chem.* **98**, 4352 (1994).

²⁵E. P. Plueddemann, *Silane Coupling Agents*, 2nd ed. (Plenum, New York, 1991).

Fabrication of hybrid superconductor–semiconductor nanostructures by integrated ultraviolet-atomic force microscope lithography

Pasqualantonio Pingue,^{a)} Marco Lazzarino, and Fabio Beltram
*Scuola Normale Superiore, I-56126 Pisa, Italy and Laboratorio Nazionale TASC-INFM,
I-34012 Trieste, Italy*

Ciro Cecconi,^{b)} Paolo Baschieri, Carlo Frediani, and Cesare Ascoli
Istituto di Biofisica del CNR, I-56127 Pisa, Italy

(Received 8 September 1996; accepted 12 May 1997)

Hybrid superconductor–semiconductor (S–Sm) nanostructures were fabricated by integrating standard ultraviolet photolithography and direct patterning of photoresist with an atomic force microscope (AFM). This novel technology was used to fabricate Nb–InAs–Nb weak links comparable in length to the coherence length. These structures exhibit high critical currents up to $10 \mu\text{A}/\mu\text{m}$ in planar geometry at 0.3 K. The fabrication protocol is based on the modification of photolithographically defined patterns by AFM *static ploughing* of the photoresist. Wet chemical etching is subsequently used for the definition of nanoscale S–Sm–S bridges. Additionally Lift-off procedures allowed the fabrication of submicron superconducting bridges. Successful fabrication of the nanostructures was verified by electrical characterization and by AFM and scanning electron microscope structural characterization. © 1997 American Vacuum Society. [S0734-211X(97)12804-9]

I. INTRODUCTION

The study of coherent transport in solids is demanding fabrication of ever smaller structures to tailor quantum confinement in all dimensions. To date, probably the most flexible fabrication approach is based on electron-beam patterning which yields structures with lateral resolution ~ 10 nm. Widespread use of this technique, however, is hindered by high investment and running costs. Alternative techniques have been suggested that are based on scanning probe microscopes¹ such as the scanning tunneling microscope (STM) and the atomic force microscope (AFM). These techniques may offer comparable or even better resolution at a fraction of the cost, but their use will probably be confined to the fabrication of test structures only due to the slower writing speed achievable.

The typical operation mode of STM lithography is analogous to e-beam patterning in that the tip is used as a source of (low-energy) electrons. AFM patterning was also demonstrated but along different lines of operation. Among these, the direct modification of the material (ploughing) was shown in the case of polymeric² or metallic³ substrates. More recently, other authors studied the use of the AFM as a tool to modify PMMA/MAA resist multilayers⁴ or ultraviolet (UV) photoresist⁵ on semiconductors.

In this article we explore the use of the AFM as a tool to extend the resolution of standard UV photolithographic processes to fabricate superconductor or hybrid superconductor–semiconductor nanostructures. These structures have attracted increasing interest lately in light of the wealth of physical phenomena that they offer.⁶ Two types of

submicron structures will be demonstrated: superconductor–semiconductor–superconductor weak links⁷ and narrow superconducting lines.

II. FABRICATION

Figure 1 shows the fabrication steps employed for the realization of weak links on epitaxially InAs doped $n = 1.5 \times 10^{18} \text{ cm}^{-3}$. Thin (~ 40 – 100 nm) niobium films were e-beam evaporated at 1×10^{-7} mbar without any surface treatment [Fig. 1(a)] and H-shaped resist (Shipley S1400-27) structures were defined by standard UV lithography [Fig. 1(b)]. The central strip is $20 \mu\text{m}$ wide and $500 \mu\text{m}$ long. The resist thickness was tuned in the range 30 – 300 nm by adjusting spinner velocity (6000 – 8000 rpm) and resist dilution (from $1:1$ to $1:15$). After hard baking for 20 min at 110°C , the AFM lithographic process was carried out. This was performed by direct writing with the AFM tip (static ploughing²) to obtain the narrow gaps desired [Fig. 1(c)]. Wet chemical etching ($50 \text{ ml H}_2\text{O}_2 + 2 \text{ ml H}_2\text{O} + 1 \text{ g NaOH}$) was used to selectively remove the Nb with negligible damage to the InAs surface.

Fabrication of the submicron-wide Nb links is schematically illustrated in Fig. 2. In this case undoped GaAs substrates were employed [Fig. 2(a)] and rectangular openings [Fig. 2(b)] were defined by standard UV lithography. Several separations in the opening were studied in the range 2 – $9 \mu\text{m}$. Direct writing by a line scan of the AFM tip was then used to connect the openings [Fig. 2(c)]. In order to optimize the lift-off process, the structure was etched with $\text{H}_3\text{PO}_4:\text{H}_2\text{O}_2:\text{H}_2\text{O}$ ($3:1:50$) to remove ~ 100 nm of the GaAs [Fig. 2(d)]. Following this etch, 100 nm of niobium were e-beam evaporated [Fig. 2(e)] followed by the lift-off process [Fig. 2(f)].

^{a)}Electronic mail: pingue@aix6sns.sns.it

^{b)}Present address: Institute of Molecular Biology, University of Oregon, Eugene, OR 97403.

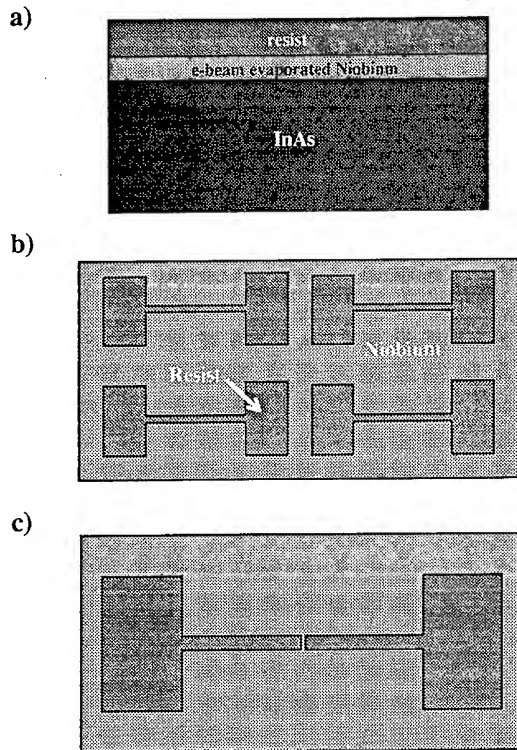


FIG. 1. Schematic description of the steps used for the fabrication of weak links. (a) Cross section of initial layers; (b) UV lithography-defined resist pattern; (c) final resist pattern after AFM static ploughing of the resist. The pattern in (c) is transferred to the Nb film by wet chemical etching (see the text).

A homemade AFM operating in air⁸ was utilized for both the fabrication and the structural characterization. Key elements determining the ultimate resolution of this technique are the shape and the lateral stability of the AFM probe. Silicon AFM probes with triangular-shaped cantilevers⁹ were chosen to minimize the lateral torsion during ploughing. We performed a series of line scans orthogonal to the resist strip until the topographical signal was flat. Cuts were made by applying a normal force of $\sim 1 \mu\text{N}$ at scan velocities between 10 and 100 $\mu\text{m/s}$. Depending on the final geometry, between 5 and 20 s were necessary to completely remove the resist. Removal of the resist is strongly promoted by tip–edge collision and is subsequently completed by the normal force applied. An interesting aspect of AFM nanofabrication is the possibility of monitoring the fabrication process by exploiting AFM imaging capabilities. It should be mentioned, however, that after several cuts are performed tip damage is often observed. This deterioration manifests itself by a poorer image quality and requires tip replacement after about 10 cuts. This same effect limits the repeatability of the resist lithography which, however, for the first cuts is very high and is essentially limited by tip positioning (negligible influence) and shape (sharper tips being, of course, preferred).

III. STRUCTURAL CHARACTERIZATION

Nanostructures were thoroughly characterized by AFM and scanning electron microscopy (SEM).

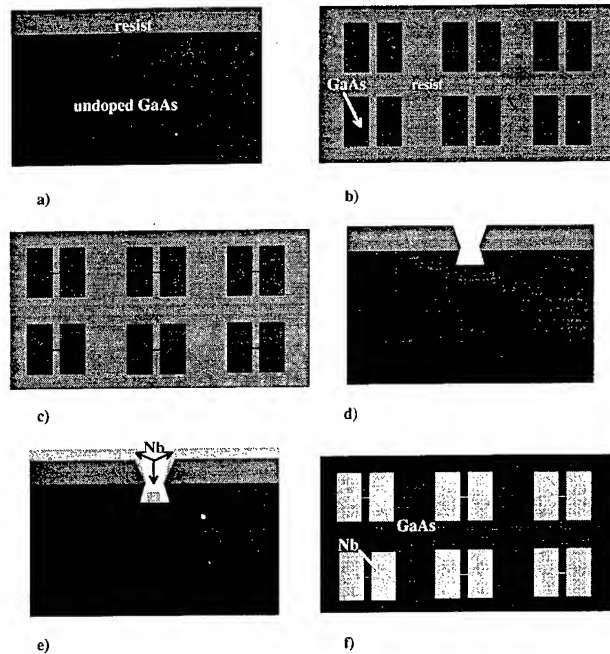


FIG. 2. Schematic description of the steps used for the fabrication of superconducting microbridges. (a) Cross section of initial layers; (b) UV lithography-defined resist pattern; (c) final resist pattern after AFM static ploughing of the resist; (d) cross section after wet chemical etching of the semiconductor; (e) Nb e-beam evaporation; (f) final Nb structures after lift-off process.

Figure 3 shows a few completed weak links among the shortest ones fabricated, at various enlargements. Figures 3(a) and 3(b) demonstrate the good uniformity achieved along the gap in the case of 100- and 40-nm-thick Nb films, respectively. Particularly in the context here of interest, a very important feature of the present technique is the capability of the AFM to unambiguously verify the complete removal of the superconducting film. This can be more convincingly obtained compared to SEM measurements due to the accuracy of depth profiling yielded by the AFM [see Figs. 3(c) and 3(d) for the case of a ~ 40 -nm-thick Nb film]. Figures 3(b) and 3(d) show that AFM lithography yields better defined resist edges than standard UV processing even after wet-chemical etching.

Gap widths as small as 250 nm were obtained, with lateral scaling being calibrated on a reference grating. Several factors contribute to limiting the shortest gap achievable, namely, tip shape, etch kinetics, and film and resist thickness. For our particular application the need for rather aggressive chemicals to remove the Nb layer poses a limitation on the resist thickness. Additional limitations stem from the need to ensure total film removal which requires longer etching times.

SEM and AFM images of samples of fabricated 100-nm-thick Nb microbridges are shown in Fig. 4 for different lengths and widths. Continuous lines were achieved with widths as small as 200 nm. We should like to emphasize that a broad range of aspect ratios can be achieved for the superconducting strip with the present approach. Thicker lines can, in fact, be easily obtained by deeper etching of the semi-

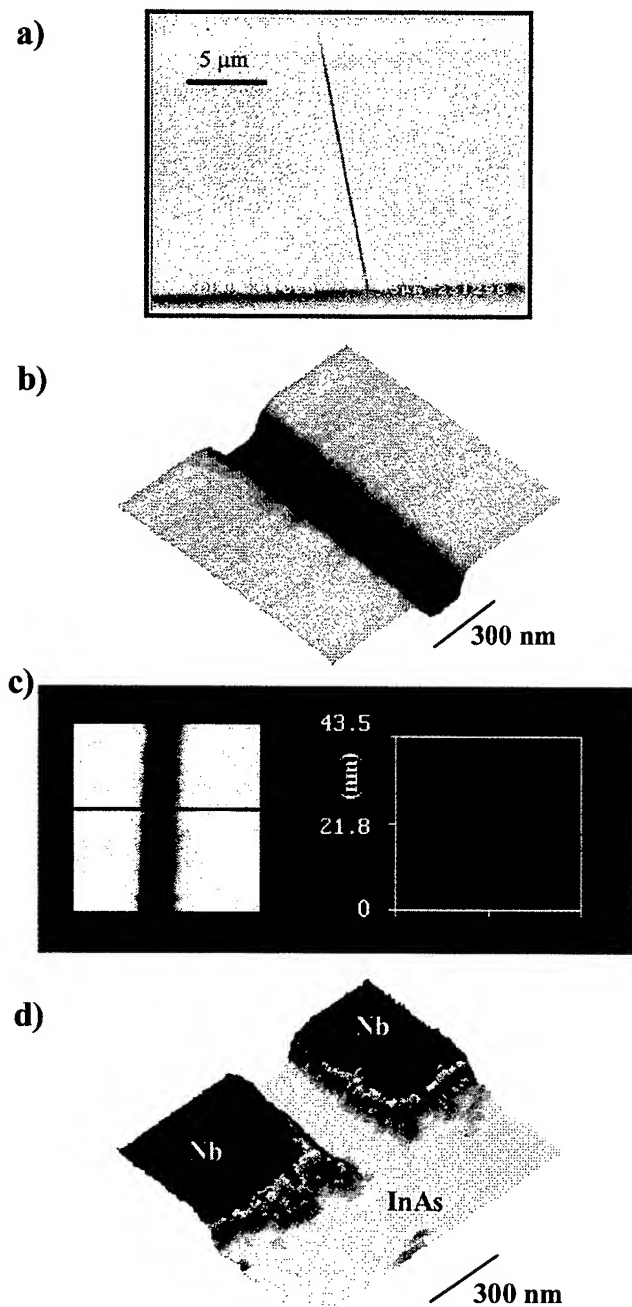


FIG. 3. Structural characterization of the weak links fabricated. (a) SEM micrograph of an entire link; (b)–(d) AFM images of sections of a link (see the text).

conductor substrate with limited effect on the final linewidth. Some of the structures exhibited a depression in the pad region along the bridge direction. This can probably be linked to unintentional ploughing of the semiconductor substrate which is greatly enhanced by the subsequent chemical etch.

IV. ELECTRICAL CHARACTERIZATION

In order to fully verify the efficacy of the fabrication procedures presented we performed low-temperature dc four-wire current–voltage (I – V) measurements in a closed-cycle ^3He cryostat.

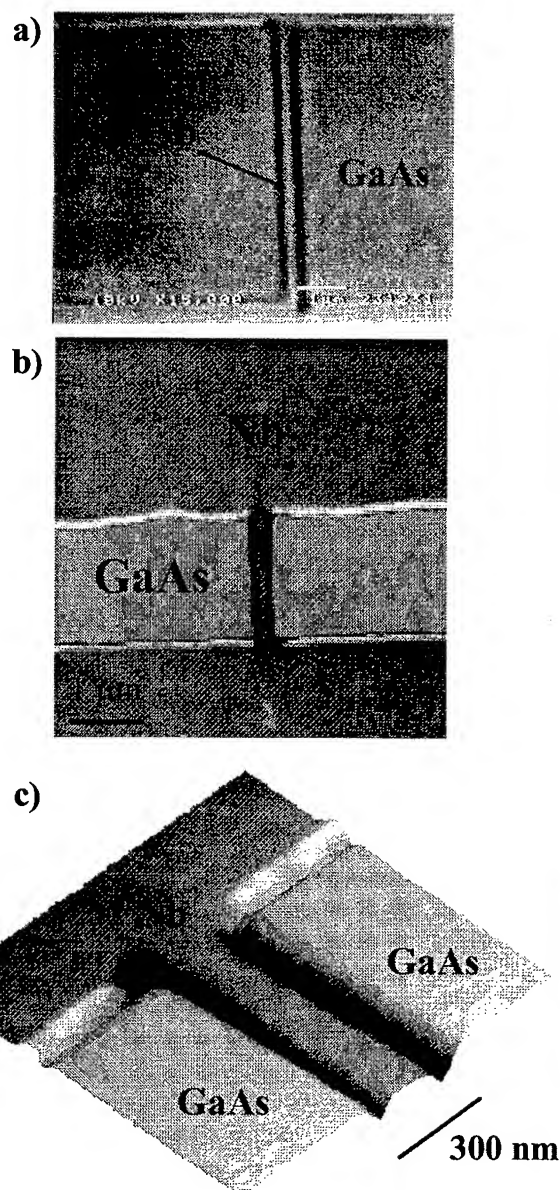


FIG. 4. Structural characterization of several microbridges. (a) SEM micrograph of a 5.6- μm -long bridge; (b)–(c) AFM images of a 1.7- μm -long bridge.

Figure 5(a) shows the I – V characteristic at 0.34 K for a 20- μm -wide and 260-nm-long Nb–InAs weak link. The coherence length in our $n = 1.5 \times 10^{18} \text{ cm}^{-3}$ InAs substrate is estimated to be $\sim 120 \text{ nm}$, a value comparable to the gap length. This guarantees significant semiconductor proximization¹⁰ and yields a measured critical current of about 200 μA . This large value allowed us to carefully study the temperature and magnetic-field dependence of the critical current in the hybrid structure.¹¹ As mentioned above, the use of the AFM ensures that no residual Nb film or debris is shorting the link thus guaranteeing the validity of the measurements.

Figure 5(b) shows the I – V characteristic at 0.34 K for our Nb microbridges (the inset shows the I – V at 6 K); the continuity of the film is proven by the high value of the critical

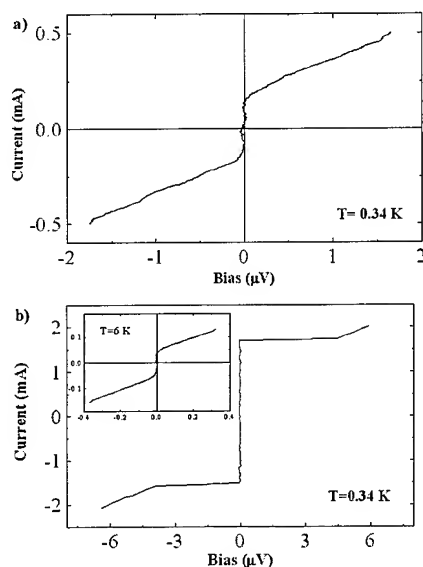


FIG. 5. I – V characterization of the devices. (a) I – V at 0.34 K of a 260-nm-long, 20- μ m-wide weak link; (b) I – V at 0.34 and 6 K (see the inset) of a 4.6- μ m-long 210-nm-wide microbridge.

current. Structures with these dimensions are of interest in several fields, e.g., in the study of vortex dynamics in type-II superconductors.¹²

V. CONCLUSIONS

In conclusion we have presented a novel implementation of AFM nanolithography that can be successfully exploited to extend the resolution limits of standard UV photolithography. Lines of the order of 100 nm were both removed and fabricated in the case of superconducting films. Successful fabrication was demonstrated by electrical characterization

of the completed devices. The present approach can find other fields of application and probably yield higher resolution in the case of planar semiconductor nanostructures, e.g., for the implementation of point contacts¹³ and nanogates.

ACKNOWLEDGMENT

Partial financial support from Consiglio Nazionale delle Ricerche within the framework of an agreement between CNR and Scuola Normale Superiore is gratefully acknowledged.

¹*Technology of Proximal Probe Lithography*, edited by C. R. K. Marrian (SPIE, Bellingham, WA, 1993), and references therein.

²T. A. Jung, A. Moser, H. J. Hug, D. Brodbeck, R. Hofer, H. R. Hidber, and U. D. Schwarz, *Ultramicroscopy* **42–44**, 1446 (1992); X. Jin and W. N. Unertl, *Appl. Phys. Lett.* **61**, 657 (1992); E. Boschung, M. Heuberger, and G. Dietler, *ibid.* **64**, 3565 (1994).

³H. Gobel and P. Von Blanckenhagen, *J. Vac. Sci. Technol. B* **13**, 1247 (1995); T. Sumonogi, T. Endo, and K. Kuwahara, *ibid.* **13**, 1257 (1995).

⁴L. L. Sohn and R. L. Willet, *Appl. Phys. Lett.* **67**, 1552 (1995).

⁵M. Wendel, S. Kühn, H. Lorenz, and J. P. Kotthaus, *Appl. Phys. Lett.* **65**, 1775 (1994).

⁶See, for example, A. W. Kleinsasser and W. J. Gallagher, in *Superconducting Devices* edited by D. Rudman and S. Ruggiero (Academic, Boston, 1990); T. M. Klapwijk, *Physica B* **197**, 481 (1994).

⁷B. D. Josephson, in *Superconductivity*, edited by R. D. Parks (Dekker, New York, 1969), Chap. 9.

⁸M. Allegrini, E. Arpa, C. Ascoli, P. Baschieri, F. Dinelli, C. Frediani, M. Labardi, A. Lio, T. Mariani, and L. Vanni, *Nuovo Cimento D* **15**, 279 (1993).

⁹Park Scientific Instruments, Ultralevers STUL-20AU AFM silicon tips. The tip radius is about 10 nm, the tip angle is 20°, and the spring constant is about 18 N m⁻¹.

¹⁰P. De Gennes, *Rev. Mod. Phys.* **36**, 225 (1964).

¹¹P. Pingue *et al.* (unpublished).

¹²P. Pingue *et al.* (unpublished).

¹³B. J. van Wees, H. van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. van der Marel, and C. T. Foxon, *Phys. Rev. Lett.* **60**, 848 (1988).

Metal-based single electron transistors

W. Chen and H. Ahmed

Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom

(Received 8 September 1996; accepted 15 February 1997)

Metal-based single electron transistors operating at liquid nitrogen temperature were fabricated. Multiple tunnel junctions with metal particles of about 3 nm in diameter were used to construct the transistors. In these transistors the gates are placed directly underneath the islands and separated only by a 10-nm-thick dielectric. The capacitance of the islands is less than 1 aF. Clear oscillations induced by gate bias voltage are observed in the current-voltage characteristics at 77 K. © 1997 American Vacuum Society. [S0734-211X(97)05804-6]

I. INTRODUCTION

Single electron charging or the Coulomb blockade effect is a phenomenon where the electrical conduction through an isolated metallic island, coupled to an external circuit through tunneling junctions, directly reflects the quantization of the electron charge on the island. Single electron charging effects enable us to observe the transfer of single electrons in a device by measuring macroscopic changes in its electrical characteristics. These effects have been studied extensively because of the possibility of making microelectronic devices based on controlling precise numbers of electrons rather than average numbers of electrons. The temperature at which these effects can be observed in a device is inversely proportional to the capacitance of the device which in turn is proportional to the size of the device. Therefore the charging effects can only be observed in very small structures with very small capacitance. The operating temperature is usually below 4 K if the structures are made entirely by lithography methods. Some researchers showed that, by employing nanoparticles¹ or nanocrystals² to form tunneling junctions, their size is of a few nanometers and is not limited by the lithography resolution; single electron charging effects can be observed at liquid nitrogen temperature.

The single electron transistor is an electronic device in which the conduction through the device is controlled by single electron charging effects and was first demonstrated by Fulton and Dolan³ using Al/Al_xO_y/Al tunnel junctions fabricated by electron beam lithography and multiple-angle deposition-oxidation-deposition cycles. The size of their device was limited by the resolution of the lithography hence the capacitance of the device is usually in the fF range making observation of single electron oscillations only possible at temperatures lower than 2 K. The operating temperature can be raised by scaling down the structures to a smaller size and work has been carried out to raise the operating temperature of single electron devices. In metallic structures, one of the approaches was to use anodization to reduce the size of Al/Al₂O₃/Al tunnel junctions and by doing so single electron devices operating at 30 K were demonstrated.⁴

The simplest single electron transistors consist of two tunnel junctions between source and drain and a gate. However, it has been shown that a multiple-tunnel junction (MTJ) which consists of a series of small islands is more suitable,

although more complicated, for single electron devices.⁵ Since the electrical characteristics do not depend on the details of each of the tunnel junctions and islands, natural metal grains of very small sizes can be used to form MTJ structures. Single electron transistors have been fabricated with such structures and oscillations were observed at liquid nitrogen temperature in side gated devices.⁶ The distance between the conduction channel and the gate in these devices was, however, very large (~100 nm) so a large gate voltage (the period of single electron oscillations was about 6 V) was needed to control the conduction. In order to construct single electron memory and logic devices a more effective gate control and a low gate voltage are needed. In this article we report the fabrication process and electrical characteristics of a metallic single electron transistor in which the gate is placed directly underneath the islands separated only by a thin layer of dielectric of a few nanometers thickness; with this more effective gate control can be achieved.

II. FABRICATION PROCESS

The metallic single electron transistors discussed in this article were fabricated on silicon substrates with a 240-nm-thick layer of thermally grown SiO₂ on top. Gold bond pads and connecting leads were formed by optical lithography and metallization. All the devices have nominally the same geometry in which the islands are separated by a thin layer of dielectric from the gate. The active area of the devices is shown schematically in Fig. 1. To reduce the step coverage problem the gate is situated in a recess etched in the substrate. The depth of the recess was chosen to be the same as the thickness of the gate metal, in our case a 20-nm-thick aluminum layer. The gate recess was fabricated by electron beam lithography with PMMA resist followed by etching in 1:1 de-ionized water:buffered HF. The etch rate is about 1 nm/s so that accurate control of etch depth can be achieved. After gate recess etching, the sample was cleaned with argon plasma etching and a layer of aluminum was deposited in a sputter coater and then lifted off in acetone. The sputter coater was used to deposit the aluminum layer to get better step coverage. Gate continuity was examined before the next step of processing.

Another step of electron beam lithography was performed to deposit a layer of gate dielectric which was aligned to the

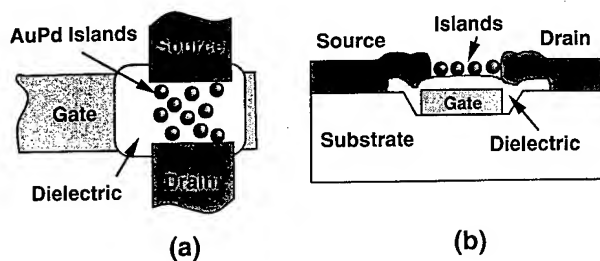


FIG. 1. Schematic of a single electron transistor. (a) A plan view of the device. (b) A cross-sectional view of the active area.

gate metal. Devices were fabricated using both SiO_2 and Si_3Ni_4 in our experiments and the difference in their electrical characteristics will be discussed later. The sputter coater was used for the deposition of dielectric layers and the deposition parameters are shown in Table I. After the deposition the samples were put in acetone for liftoff and islands were formed in the same way as that described earlier.⁶ Finally source-drain contacts were connected to the islands. Due to the step coverage problem, two steps of metallization were used. A thin layer of NiCr/Au about 10 nm in thickness was used to define the source-drain gap and a thicker layer of NiCr/Au of 40 nm was used to cover the steps. Figure 2 shows one of the completed devices. The width of the leads is about 30 nm and the gap between source and drain contacts is also about 30 nm. The distance between the aluminum gate and the islands is the total thickness of natural aluminum oxide plus the dielectric thickness and is about 10 nm. In some devices no dielectric was deposited on the gate metal and the islands were deposited directly onto the aluminum gate. It was found that the natural oxide on the aluminum gate was, however, too thin to form an effective barrier and the gate leakage current measured was too great.

III. ELECTRICAL CHARACTERISTICS

The completed devices were diced and bonded into ceramic packages that were then mounted into a dip stick for low temperature measurements. The I - V characteristics of the devices were measured at room temperature and at liquid nitrogen temperature using Keithley 236 source measuring units (SMUs). One SMU was used as the source for a voltage across the source and drain electrodes and to measure the current through the device; a second SMU was used to apply a voltage to the gate electrode. For low temperature measurements the dip stick was immersed in a liquid nitrogen Dewar.

At room temperature all working devices showed a nonlinearity in the source-drain voltage-current characteristics when the gate electrode was kept at ground potential. No

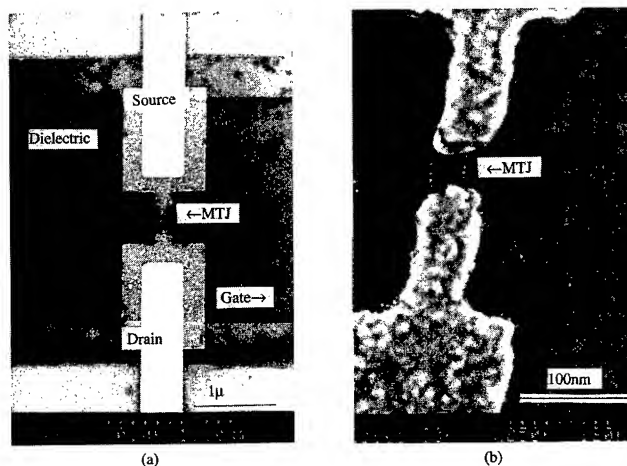


FIG. 2. SEM micrographs of a completed device. (a) Low magnification micrograph in which different stages of metallization can be seen. (b) High magnification micrograph showing details of the multiple tunneling junctions. The islands are about 3 nm in diameter.

periodic oscillations could be detected in the source-drain conductance when the gate voltage was swept with the source-drain voltage kept constant because of thermal smear out. At liquid nitrogen temperature (77 K) a pronounced nonlinearity was observed in the source-drain I - V characteristics and in some devices clear Coulomb gaps were seen. Oscillations appeared in the source-drain current when the gate voltage was swept and the source-drain voltage was kept constant. There was no oscillation in the gate current measured at the same time. In Fig. 3 the source-drain current in a device with a Si_3Ni_4 gate dielectric is shown as a function of gate bias voltage which was swept from -2 to 2 V. The data were taken at 77 K and different curves were measured at different source-drain bias voltages from 0 to 250 mV in 50 mV steps. As can be seen in Fig. 3, several distinct large peaks appear in the curves and within each large peak there are several small peaks. In a double junction single electron transistor the source-drain current oscillates with a

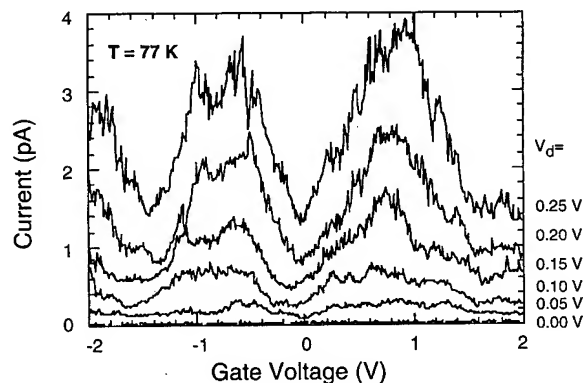


FIG. 3. Source-drain current as a function of gate voltage at a constant source-drain bias voltage V_d measured at 77 K for a device with a Si_3Ni_4 dielectric. Different curves show current measured at different V_d from 0 to 250 mV in 50 mV steps.

TABLE I. Deposition parameters of gate dielectric layers.

Dielectric	Flow rate (sccm)	Chamber pressure (mB)	Deposition rate (nm/min)	Power (W)	Thickness (nm)
SO_2	3.6	5×10^{-4}	7.9	50	10
Si_3Ni_4	3.6	5×10^{-4}	3.3	50	10

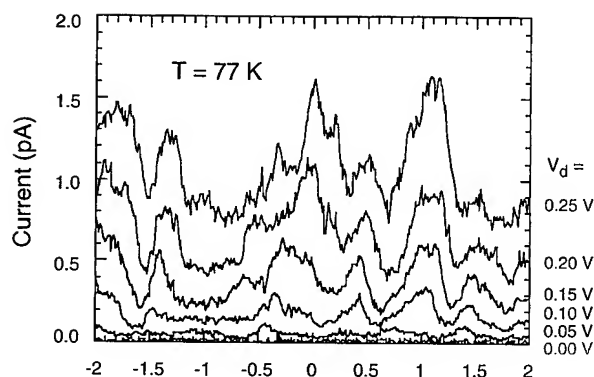


FIG. 4. Source-drain current as a function of gate voltage at a constant source-drain bias voltage V_d measured at 77 K for a device with a SiO_2 dielectric. Different curves show current measured at different V_d from 0 to 250 mV in 50 mV steps.

period of e/C_g when its gate voltage is swept; C_g is the gate capacitance. It was shown that when MTJs instead of double junctions are used in single electron transistors the oscillations in the source-drain current will have multiple periodicity even in the case of MTJs consisting of identical islands and junctions.⁵ Here the broadest period does correspond to the gating effect and, taking the broadest period in Fig. 3, the gate capacitance is estimated to be 0.13 aF from $e/C_g \approx 1.5$ V. There are many peaks in the Fourier transformed data, the most prominent one being at 1.6 V and the second most prominent at 3.1 V. Because in our experiment the transistors consist of multiple tunneling junctions and the size of the islands used to form the MTJs is not uniform, the oscillations are very complicated. As shown by Nakazato *et al.*⁵ multiple periodicity will develop for single electron transistors made of MTJs. In our device the additional complexity is the two-dimensional nature of the tunneling junctions. Not only is there charging up of different islands in the conduction path but there are also those that are not in the conduction path but due to their proximity to it will contribute to the oscillation of the source-drain current. Calculations showed that for MTJs made from identical junctions the narrow peaks in the source-drain current disappear at high drain bias.⁵ In our devices, however, this characteristic is somewhat different and new peaks developed at high source-drain bias. This can be attributed to the higher probability of electrons tunneling into islands near the conduction path. The oscillations observed in the present devices have a much shorter period than that observed in side gated devices.⁶ The period for a device with a side gate located 100 nm from the islands was about 6.3 V. The period of oscillation of the back gated device is about 1.5 V. Similar oscillations were observed in devices fabricated with a SiO_2 dielectric (Fig. 4). As more effective gate control is achieved, single electron memory or logic devices can be constructed by the fabrication technique described in this article.

In the present devices the main features of the measurement results normally remain the same but there are minor differences from different sweeps for the same devices. Characteristics are usually different for nominally identical

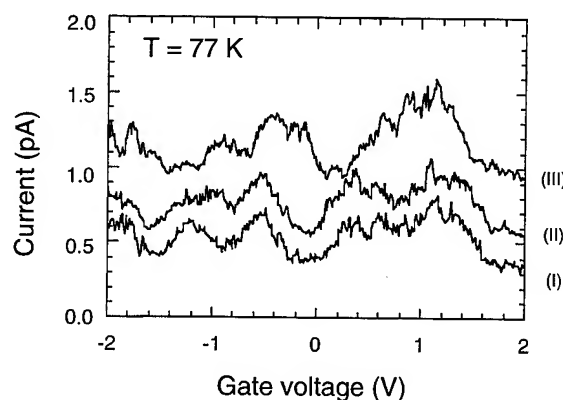


FIG. 5. Comparison of I - V characteristics of different sweeps of gate voltage for a device with a Si_3Ni_4 dielectric measured at 77 K. Curves (II) and (III) are shifted for 0.2 and 0.6 pA, respectively, for clarity. The gate voltage was swept from -2 to 2 V for curves (I) and (II) and from 2 to -2 V for curve (III).

samples. Figure 5 compares I - V characteristics of three measurements from the same Si_3Ni_4 device. Curves (II) and (III) are shifted by 0.2 and 0.6 pA, respectively, for clarity. Curves (I) and (II) were measured with gate sweeping from -2 to 2 V and they are almost identical. Curve (III) was measured with gate sweeping from 2 to -2 V and the peaks and troughs are shifted compared to those in curves (I) and (II). The reason for this shift is not yet understood. For SiO_2 devices, the measurement results are more repeatable as shown in Fig. 6. Curves (II) and (III) are shifted by 0.2 and 0.4 pA, respectively, for clarity. Gate voltages were swept from -2 to 2 V for curves (I) and (II) and from 2 to -2 V for curve (III). Unlike results from the Si_3Ni_4 devices, the curves are almost identical for both gate-voltage sweep directions. More repeatable results were obtained with a SiO_2 dielectric layer. Examination in a scanning electron microscope showed that the SiO_2 films were more uniform than the Si_3Ni_4 films.

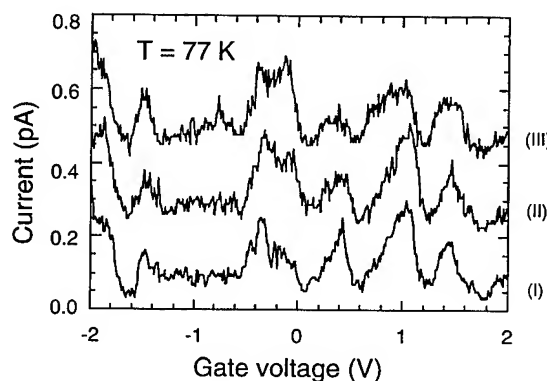


FIG. 6. Comparison of I - V characteristics of different sweeps of gate voltage for a device with a SiO_2 dielectric measured at 77 K. Curves (II) and (III) are shifted for 0.2 and 0.4 pA, respectively, for clarity. The gate voltage was swept from -2 to 2 V for curves (I) and (II) and from 2 to -2 V for curve (III).

IV. CONCLUSIONS

Metal-based single electron transistors were fabricated with multiple tunnel junctions formed in metallic granular films deposited by an ionized beam deposition technique. The islands were only about 3 nm so the devices can be operational at liquid nitrogen temperature. The gate electrodes were located directly underneath the multiple tunnel junctions and separated by a dielectric layer only a few nanometers thickness. Electrical characteristics measured at liquid nitrogen temperature showed clear single electron charging effects. The source-drain current oscillated at constant source-drain bias voltage when the gate voltage was swept. It is believed that the complicated structure in the $I-V$ characteristics was due to the charging up of islands in as well as near the conduction path. Due to the stochastic nature of the islands used in our devices the electrical characteristics from different devices are not repeatable. In comparing the devices fabricated with SiO_2 and Si_3Ni_4 dielectrics it was concluded that the SiO_2 layer has the better dielectric

properties in our present experiment. As a moderate gate voltage was used to control the conduction, it is possible to use the metal-based single electron transistors described in this article to construct single electron memory or logic devices.

ACKNOWLEDGMENT

One of the authors (W.C.) was supported by Wolfson College of the University of Cambridge with a Hitachi-Wolfson fellowship.

¹W. Chen, H. Ahmed, and K. Nakazato, *Appl. Phys. Lett.* **66**, 3383 (1995).

²D. L. Klein, J. E. B. Katari, R. Roth, A. P. Alivisatos, and P. L. McEuen, *Appl. Phys. Lett.* **68**, 2574 (1995).

³T. A. Fulton and G. J. Dolan, *Phys. Rev. Lett.* **59**, 109 (1987).

⁴Y. Nakamura, D. L. Klein, and J. S. Tsai, *Appl. Phys. Lett.* **68**, 275 (1996).

⁵K. Nakazato, R. J. Blaikie, and H. Ahmed, *J. Appl. Phys.* **75**, 5123 (1994).

⁶W. Chen and H. Ahmed, *J. Vac. Sci. Technol. B* **13**, 2883 (1995).

Electron beam dot lithography for nanometer-scale tunnel junctions using a double-layered inorganic resist

Satoshi Haraichi,^{a)} Toshimi Wada, Sucheta M. Gorwadkar,^{b)} Kenichi Ishii, and Hiroshi Hiroshima

Electrotechnical Laboratory, Tsukuba-shi, Ibaraki 305, Japan

(Received 12 September 1996; accepted 30 January 1997)

We have developed a metal-insulator-metal (MIM) junction fabrication process using top two layers of a separated by implanted oxygen substrate as a double-layered electron beam resist and a dot-exposure method in order to simplify the process sequence and fabricate smaller MIM junctions than those made by the conventional multiple-angle deposition-oxidation-deposition method. The diameter of the dot openings of a suspended mask tends to saturate with an increase in the exposed dose and is independent of the dot interval. We were able to successfully fabricate a series of Ti dots with minimum diameters of about 20 nm and much smaller junctions than the minimum dot size of 20 nm by an inorganic resist lift-off process. A one-dimensional array of six MIM junctions fabricated by this method showed typical current-voltage characteristics of tunnel junctions at room temperature. © 1997 American Vacuum Society. [S0734-211X(97)10804-6]

I. INTRODUCTION

Single electron devices are considered to be powerful candidates for future devices that overcome the integration limit in conventional Si devices.¹ Nanofabrication techniques for ultrasmall tunnel junctions are indispensable to the development of single electron devices that operate at moderate temperatures higher than the liquid nitrogen temperature. Nanometer-scale tunnel junctions have been realized by various structures such as split-gate heterostructures,² tunnel junctions in a very thin poly-Si layer,³ and metal-insulator-metal (MIM) tunnel junctions.^{4,5}

MIM small-area tunnel junctions have been realized by using the multiple-angle deposition-oxidation-deposition method^{6,7} for which suspended masks made of such organic resists as polymethylmethacrylate (PMMA) are patterned by electron beam (EB) lithography. EB exposure processes are the most promising candidates for nanofabrication technology from the point of view of patterning accuracy and throughput. However, the resolution of organic EB resist is known to be no better than around 10 nm,⁸ so a high resolution resist must be developed for nanometer-scale EB lithography. Allee and Broers demonstrated that much finer patterns can be fabricated using an inorganic resist of SiO₂ thin layers, utilizing the enhanced etch rate of EB-exposed SiO₂ in a solution.⁹ Hiroshima *et al.* demonstrated the efficiency of this method by delineating line patterns with widths of about 5 nm in a SiO₂ thin layer with a 3 nm probe.¹⁰ Recently Wada *et al.* developed a SiO₂/poly-Si double-layered EB resist process for suspended mask fabrication and applied it to Al/Al₂O₃/Al tunnel junction fabrication by using the multiple-angle deposition-oxidation-deposition method.¹¹ These fabricated tunnel junctions exhibited a clear Coulomb staircase at 12 K. There were,

however, several problems, such as irregularities or edge roughing of the mask patterns due to the granular structure of poly-Si and the nonuniform electrical characteristics of the MIM junctions that was probably due to the granular deposition of Al. To overcome process problems such as these, we have introduced a SiO₂/c-Si double-layered EB resist process originating from a separation by implanted oxygen (SIMOX) substrate and selected Ti as a metal for the junctions. Gorwadkar *et al.* have reported improvement in the basic inorganic double-layered EB resist process.¹²

In this study, we have developed another MIM junction fabrication process using top two layers of a SIMOX substrate as a double-layered EB resist and a dot-exposure method in order to simplify the process sequence and fabricate smaller MIM junctions than those made by the conventional multiple-angle deposition-oxidation-deposition method. We were able to successfully fabricate a series of Ti dots with minimum diameters about 20 nm and much smaller junctions than the dot size of 20 nm by an inorganic resist lift-off process. A one-dimensional (1D) array of six MIM junctions fabricated by this method showed typical current-voltage (*I-V*) characteristics of tunnel junctions at room temperature.

II. EXPERIMENT

The schematic sequence of a double-layered inorganic EB resist process using a dot-exposure method for suspended mask fabrication, metal dot deposition, and inorganic resist liftoff is shown in Fig. 1. The SIMOX wafers were subjected to subsequent stages of thermal oxidation and chemical etching; the samples with resist layers of 55 nm SiO₂ and 70 nm c-Si were prepared on 116 nm SiO₂/Si(100) substrates. The double-layered inorganic resist was selectively irradiated with a 50 kV finely focused EB with spot doses ranging from 2.5 to 25 pC/dot [Fig. 1(1)]. After EB irradiation, the samples were processed in a barrel-type oxygen plasma asher for 15 min to remove the contamination layer due to the EB

^{a)}Electronic mail: haraichi@etl.go.jp

^{b)}Currently with the New Energy and Industrial Technology Development Organization (NEDO), Japan.

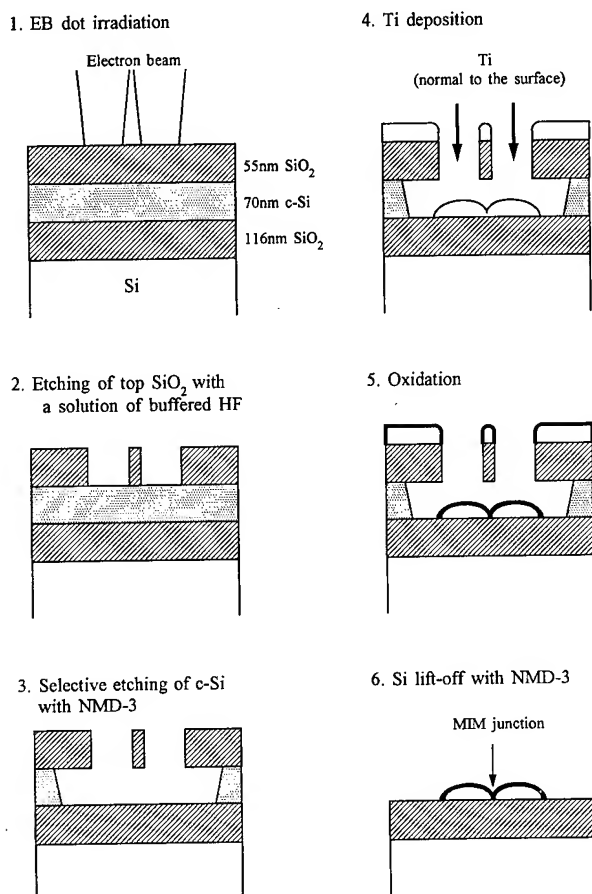


FIG. 1. Schematic of the sequence of a double-layered inorganic EB resist process using a dot-exposure method for suspended mask fabrication, metal dot deposition, and inorganic resist lift-off.

irradiation. The oxygen pressure was about 1 Torr and the rf power was 100 W. The top SiO_2 layer was then etched with a solution of buffered HF ($\text{HF}:\text{NH}_4\text{F}:\text{H}_2\text{O}=1:5:40$) at room temperature [Fig. 1(2)]. The etching rate of the EB-exposed SiO_2 was enhanced in buffered HF (BHF), and the selectivity (the ratio of the etch rate of the exposed region to that of the unexposed region of the top SiO_2 layer) of the EB area dose of $1 \text{ C}/\text{cm}^2$ was estimated to be 3.3 according to our previous work.¹¹ The underlying c-Si layer was etched by a commercially available photoresist developer, NMD-3 solution (Tokyo Ohka Kogyo Co., Ltd., Japan) at 50°C [Fig. 1(3)]. The main chemical component of NMD-3 is tetramethyl ammonium hydroxide (TMAH). We got good undercut structures easily because of the high selectivity (the ratio of the etch rate of Si(111) to that of SiO_2) of the NMD-3, about 1000 at 50°C .¹² EB evaporated Ti was vertically deposited onto a sample across a series of dot openings at a deposition rate of about 0.2 nm/s [Fig. 1(4)]. The thickness of the deposited Ti film was about 10 nm. The surface of the deposited Ti film was oxidized by exposing the sample to air, and a saturated oxidation layer of TiO_x with a thickness of about 3–4 nm (Ref. 13) was formed [Fig. 1(5)]. As it is extremely difficult to lift off inorganic resist layers completely, we lifted off only a small part of the resist layers, about $50 \mu\text{m} \times 50 \mu\text{m}$,

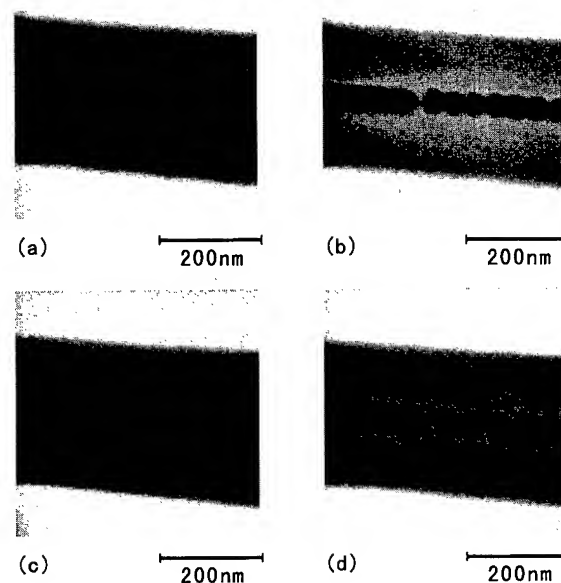


FIG. 2. SEM micrographs showing the suspended masks made by the dot-exposure method for several exposure conditions where (a) exposed dose (D) is 5 pC/dot, dot interval (DI) is 40 nm; (b) $D=10 \text{ pC/dot}$, $\text{DI}=40 \text{ nm}$; (c) $D=5 \text{ pC/dot}$, $\text{DI}=50 \text{ nm}$; and (d) $D=10 \text{ pC/dot}$, $\text{DI}=50 \text{ nm}$.

around the junction pattern which was surrounded by one or two successive frames of grooves.¹² The top Ti/SiO_2 layers inside these frames of grooves were lifted off by the underlying c-Si layer etched with the NMD-3 solution at 70°C for about 60 min under intermittent ultrasonic agitation [Fig. 1(6)]. Au/Cr pad patterns were then made by a PMMA lift-off process, and the I – V characteristics of a 1D array of MIM junctions were measured.

III. RESULTS AND DISCUSSION

Figures 2(a)–2(d) are the scanning electron microscopy (SEM) micrographs of the suspended masks fabricated by the dot-exposure method and the two step development using BHF and NMD-3 solutions for several exposure conditions. Dark gray regions correspond to openings of the top SiO_2 mask layer, and the surrounding gray regions correspond to undercuts of the c-Si layer etched along the (111) crystal plane.¹² The diameter of the dot openings of a suspended mask tends to increase with an increase of the exposed dose and shows a small dependence on the dot interval (DI). The exposed dose dependence of the diameter of the dot openings for several DIs of 50, 40, and 30 nm are shown in Fig. 3. The diameter of the dot openings tends to saturate with an increase in the exposed dose and is almost independent of the dot interval. The SiO_2 of the top resist layer is not affected very much by the proximity effect because of the much higher threshold energy of bond breaking in the resist material than in usual organic resists. With our process we can design the diameter and the interval of the dot openings independently in order to fabricate a series of dot mask openings for MIM junctions with DIs ranging more than 30 nm.

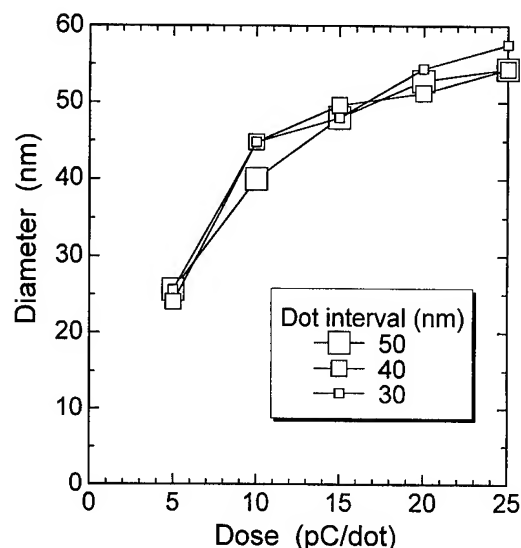


FIG. 3. The dependence of the diameter of the dot openings of a suspended mask on exposed dose for dot intervals of 50, 40, and 30 nm.

We studied the temperature dependence of the etching rate of c-Si along the (111) crystal plane in NMD-3 solution, and this is shown in Fig. 4(a). The etching rate tends to rise with an increase in the temperature, and a sufficiently high etching rate of about 25 nm/min for an inorganic resist lift-off process is obtained at 70 °C. We studied the effect of NMD-3 solution on the sheet resistance of the Ti film at 70 °C. The sheet resistance measurements, shown in Fig. 4(b), were carried out using the four-probe method to estimate the change in thickness of the Ti film due to the

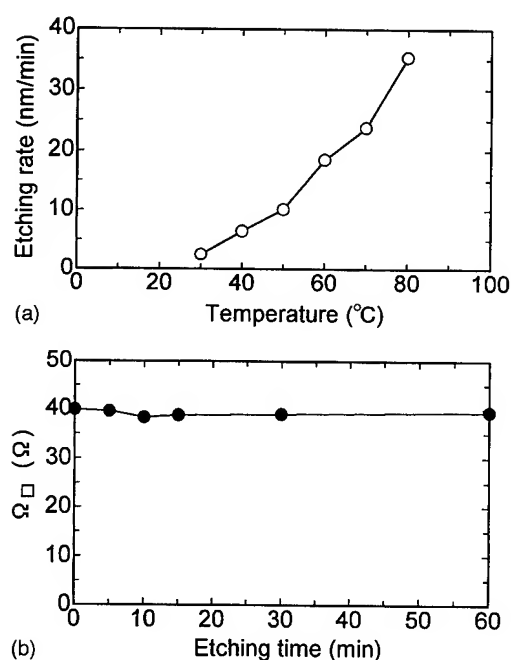


FIG. 4. (a) Temperature dependence of the etching rate of c-Si in NMD-3 solution and (b) etching time dependence of the sheet resistance of Ti films in NMD-3 solution at 70 °C.

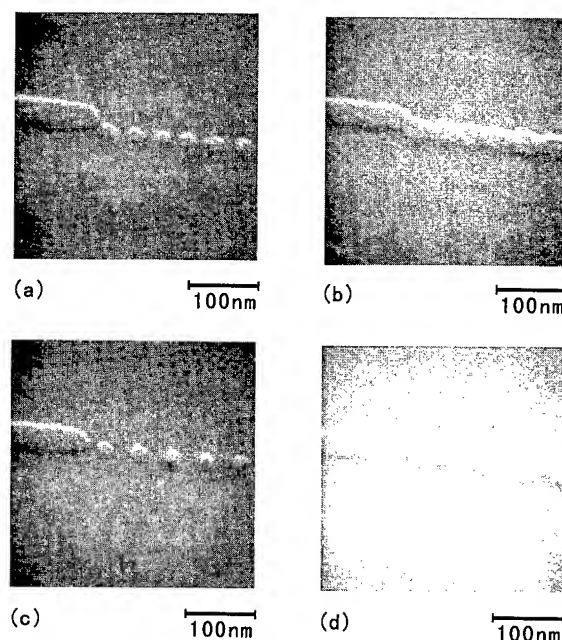


FIG. 5. SEM micrographs showing the series of Ti dots fabricated by an inorganic resist lift-off process in NMD-3 solution for several exposure conditions where (a) exposed dose (D) is 5 pC/dot, dot interval (DI) is 40 nm; (b) $D = 10$ pC/dot, DI=40 nm; (c) $D = 5$ pC/dot, DI=50 nm; and (d) $D = 10$ pC/dot, DI=50 nm.

NMD-3 solution; the film had an initial thickness of about 15 nm. The sheet resistance of the Ti film did not effectively change after treatment with NMD-3. This indicates that the Ti film is strong enough not to be affected by the NMD-3 solution because of the native oxide layer that developed on the surface. Passivation of metallic Ti by its native oxide, which is about 4 nm thick, provides a very corrosion resistant material.^{13,14}

Figures 5(a)–5(d) are SEM micrographs of the series of Ti dots fabricated by an inorganic resist lift-off process in NMD-3 solution at 70 °C using the masks shown in Figs. 2(a)–2(d), made by different EB exposure conditions. We were able to successfully fabricate a series of Ti dots with minimum diameters of about 20 nm, as shown in Figs. 5(a) and 5(c). In Fig. 5(d), we can see very small neck regions of less than 10 nm in width between the Ti dots that could be much smaller junctions than the minimum dot size of 20 nm. The exposed dose dependence of the diameter of fabricated Ti dots for several selected dot intervals of 50, 40, and 30 nm is shown in Fig. 6. The fabricated Ti dots always have diameters about 5 nm smaller than those of corresponding mask openings, shown in Fig. 3. The mask openings have sloped side walls due to the relatively low selectivity of SiO_2 of 3.3, and the diameters of the fabricated Ti dots might be almost equal to the bottom diameters of the corresponding mask openings. There seems to be very little effect of surface diffusion of Ti atoms on metal Ti deposition at room temperature.

A SEM micrograph of a 1D array pattern of the suspended mask for six MIM junction fabrication is shown in Fig. 7. Dark gray lines and dots correspond to openings of

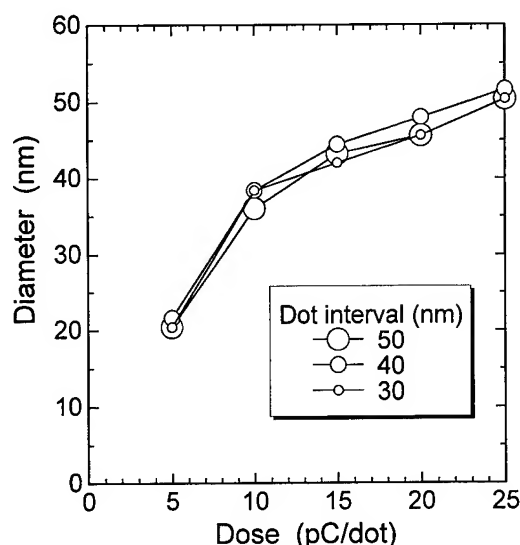


FIG. 6. The dependence of the diameter of fabricated Ti dots on exposed dose for dot intervals of 50, 40, and 30 nm.

the top SiO_2 mask layer, and the surrounding triangularly shape gray regions correspond to undercuts of the c-Si layer etched by NMD-3 solution. The series of five dots makes six MIM junctions between the lines and dots. The two horizontal line patterns connected with Au/Cr pad patterns and the I - V characteristics of the 1D array of the six MIM junctions were measured. The vertical line pattern is a gate pattern but it was not used in this study. Figure 8 shows the I - V characteristics at room temperature of a 1D array of six MIM junctions fabricated by the inorganic masks that was made

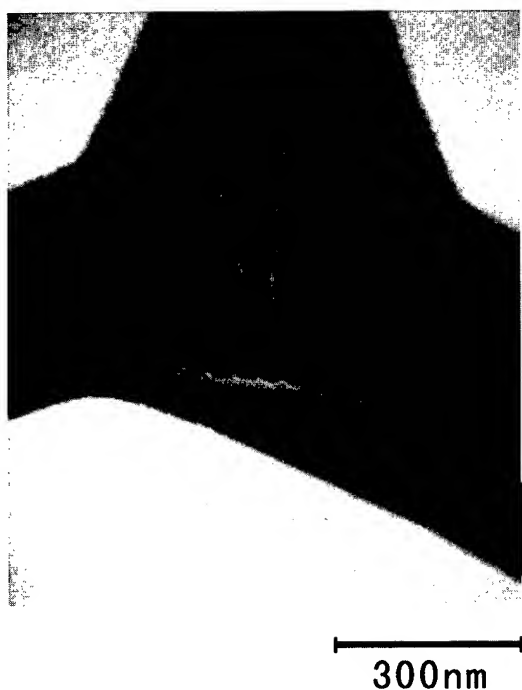


FIG. 7. SEM micrograph showing a 1D array pattern of the suspended mask for six MIM junctions fabrication.

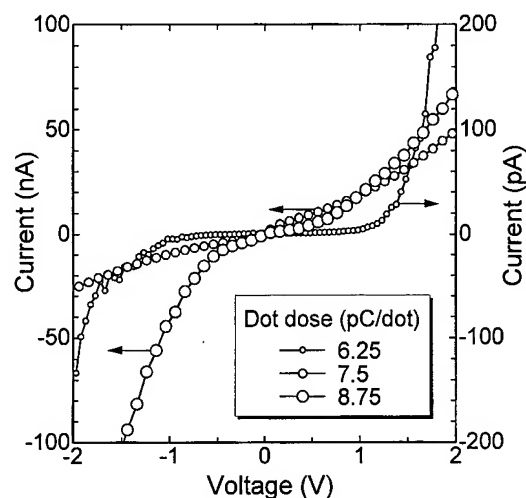


FIG. 8. I - V characteristics at room temperature of 1D array of six MIM junctions fabricated by the inorganic masks that were made by exposed doses of 6.25, 7.5, and 8.75 pC/dot.

by several exposed dose conditions, 6.25, 7.5, and 8.75 pC/dot. The dot intervals were designed equally to 40 nm. The I - V characteristics, corresponding to exposed doses of 7.5 and 8.75 pC/dot, show a large leakage current, and those junctions are almost like metal wires. The I - V characteristics, corresponding to an exposed dose of 6.25 pC/dot, show the typical I - V curve of tunnel junctions. We obtained a barrier height of 1.85 eV and a barrier thickness of 1.46 nm using a simple theory for determination of the potential barrier shape in tunnel junctions.¹⁵ We fabricated a good tunnel barrier of TiO_2 .

IV. CONCLUSIONS

We have developed a MIM junction fabrication process using the top two layers of a SIMOX substrate as a double-layered EB resist and a dot-exposure method in order to simplify the process sequence and fabricate smaller MIM junctions than those made by the conventional multiple-angle deposition-oxidation-deposition method. The diameter of the dot openings of a suspended mask tends to saturate with an increase in the exposed dose and is independent of the dot interval. The SiO_2 of the top resist layer is not affected very much by the proximity effect because of the much higher threshold energy of bond breaking in the resist material than in usual organic resists. We were able to successfully fabricate a series of Ti dots with minimum diameters of about 20 nm and much smaller junctions than the minimum dot size of 20 nm by using an inorganic resist lift-off process in NMD-3 solution at 70 °C. The diameters of fabricated Ti dots are almost equal to those of corresponding mask openings and there seems to be very little effect of surface diffusion of Ti atoms on metal Ti deposition. A one-dimensional array of six MIM junctions fabricated by this method showed typical current-voltage characteristics of tunnel junctions at room temperature.

ACKNOWLEDGMENT

This work is supported by a joint research project with the New Energy and Industrial Technology Development Organization (NEDO), Japan.

- ¹D. V. Averin and K. K. Likharev, *Single Charge Tunneling*, edited by H. Grabert and M. H. Devoret (Plenum, New York, 1992), p. 311.
- ²L. P. Kouwenhoven, A. T. Johnson, N. C. van der Vaart, C. J. P. M. Harmans, and C. T. Foxon, *Phys. Rev. Lett.* **67**, 1626 (1991).
- ³K. Yano, T. Ishii, T. Hashimoto, T. Kobayashi, F. Murai, and K. Seki, *Proceedings of the IEEE International Electron Device Meeting*, 1993, p. 541.
- ⁴P. Delsing, T. Claeson, K. K. Likharev, and L. S. Kuzmin, *Phys. Rev. B* **42**, 7439 (1990).
- ⁵H. Pothier, P. Lafarge, C. Urbina, D. Esteve, and M. H. Devoret, *Europhys. Lett.* **17**, 249 (1992).

- ⁶G. J. Dolan, *Appl. Phys. Lett.* **31**, 337 (1977).
- ⁷L. S. Kuzmin, P. Delsing, T. Claeson, and K. K. Likharev, *Phys. Rev. Lett.* **62**, 2539 (1989).
- ⁸A. N. Broers, J. M. E. Harper, and W. W. Molzen, *Appl. Phys. Lett.* **33**, 392 (1978).
- ⁹D. R. Allee and A. N. Broers, *Appl. Phys. Lett.* **57**, 2271 (1990).
- ¹⁰H. Hiroshima, S. Okayama, M. Ogura, M. Komuro, H. Nakazawa, Y. Nakagawa, K. Ohi, and K. Tanaka, *J. Vac. Sci. Technol. B* **13**, 2514 (1995).
- ¹¹T. Wada, M. Hirayama, S. Haraichi, K. Ishii, H. Hiroshima, and M. Komuro, *Jpn. J. Appl. Phys.* **34**, 6961 (1995).
- ¹²S. M. Gorwadkar, T. Wada, S. Haraichi, H. Hiroshima, K. Ishii, and M. Komuro, *Jpn. J. Appl. Phys.* **35**, 6673 (1996).
- ¹³T. Smith, *Surf. Sci.* **38**, 292 (1973).
- ¹⁴M. Jobin, M. Taborelli, and P. Descouts, *J. Appl. Phys.* **77**, 5149 (1995).
- ¹⁵J. G. Simmons and G. J. Unterkofer, *J. Appl. Phys.* **34**, 1828 (1963).

Negative differential resistance on single electron transport in a junction array of ultrasmall islands

Hiroshi Nakashima and Kiyohiko Uozumi^{a)}

Department of Physics, College of Science and Technology, Aoyama Gakuin University, Setagaya, Tokyo 157, Japan

(Received 8 September 1996; accepted 22 April 1997)

We predict a new type of negative differential resistance (NDR) in a nonlinear ten-junction array of nine conducting islands with ultrasmall self-capacitance. A main feature of our array is that two of the ten junctions go in the direction opposite to an external in-plane electric field. The NDR characteristic is one of Coulomb blockade phenomena and is triggered by competition between the two tunnel rates across two specific junctions depending on the external voltage. Calculations for the electron dynamics are based on the orthodox semiclassical single electron tunneling model.

© 1997 American Vacuum Society. [S0734-211X(97)10304-3]

I. INTRODUCTION

Some multijunction arrays with ultrasmall capacitance show steplike current-voltage (I - V) characteristics which are ascribed to single electron charging effects^{1,2} and can serve as single electron transistors.³ Recently we have theoretically shown that a nonlinear seven-junction array of six spherical conducting islands with ultrasmall self-capacitance driven by a voltage source can also show negative differential resistance (NDR) in the I - V curve due to single electron charging effects.⁴ In contrast to usual multidimensional junction arrays,^{5,6} the electrical potential is no longer concentrated within the junctions. The long-range interaction between islands was taken into account in our determinations of the island potential using the method of multiple images. An interesting feature of our seven-junction array is that the islands are arranged so that only one junction goes in the direction opposite that of an external in-plane electric field. However, in reality, the current around the NDR peak in such a seven-junction array seems to be too small for NDR to be identified clearly.

In this article, we propose a longer nonlinear ten-junction array of nine ultrasmall spherical conducting islands that shows NDR characteristics with much larger current. The new array is arranged so that two consecutive junctions go in the direction opposite that of an external in-plane electric field. The electron transport mechanisms of our seven- and ten-junction arrays that are responsible for NDR are basically the same.

To simulate stochastic electron dynamics we use the orthodox semiclassical single electron tunneling model for multijunction circuits.⁷ Let us consider arbitrary one-dimensional chains of N islands driven by a voltage V . The electron tunneling rates across the j th junction are given by a golden rule:

$$\Gamma(V, \mathbf{n} \rightarrow \mathbf{n}') = \frac{1}{e^2 R_j} \frac{\Delta E}{1 - \exp(\Delta E / k_B T)}, \quad (1)$$

where e is the electronic charge, R_j is the tunnel resistance

of the j th junction, T is the temperature, and ΔE is the electrostatic energy difference before and after the tunneling event which causes a change in the electron configuration from \mathbf{n} to \mathbf{n}' ($\mathbf{n} \equiv \{n_1, \dots, n_N\}$ is the number of excess electrons on the islands). The ΔE contains contributions from a change in the internal charging energy $\epsilon(\mathbf{n})$ as well as from the work $\kappa_j e V$ ($\sum_j \kappa_j = 1$) done by the voltage source after an electron has tunneled across the j th junction:

$$\Delta E = \kappa_j e V + \epsilon(\mathbf{n}) - \epsilon(\mathbf{n}'), \quad (2)$$

and the internal charging energy is given by

$$\epsilon(\mathbf{n}) = \frac{1}{2} e \mathbf{n} \mathbf{U}, \quad (3)$$

where \mathbf{U} is the island potential corresponding to the electron configuration. In our calculations, \mathbf{U} is estimated by using the method of multiple image charges, assuming that all the islands are spherical. If one knows the probability to find the electron configuration \mathbf{n} , $\sigma(\mathbf{n})$, then the average current is given by the net flow through any junction of the array:

$$I(V) = e \sum_{\mathbf{n}} \sigma(\mathbf{n}) [\vec{\Gamma}_j(\mathbf{n}) - \vec{\Gamma}_j(\mathbf{n}')], \quad (4)$$

where the \rightarrow/\leftarrow refers to the forward/backward tunneling across the j th junction, and the average number of excess electrons on the i th island is given by

$$\langle n_i \rangle = \sum_{\mathbf{n}} \sigma(\mathbf{n}) \mathbf{n} \cdot \mathbf{u}_i, \quad (5)$$

where \mathbf{u}_i is a unit vector in the direction of the i th island. The probability $\sigma(\mathbf{n})$ is obtained by making use of the requirement for a steady state:

$$\begin{aligned} & \sigma(n_1, \dots, n_j, \dots, n_N) [\vec{\Gamma}_j(n_1, \dots, n_j, \dots, n_N) \\ & + \vec{\Gamma}_{j+1}(n_1, \dots, n_j, \dots, n_N)] \\ & = \sigma(n_1, \dots, n_j + 1, \dots, n_N) [\vec{\Gamma}_j(n_1, \dots, n_j + 1, \dots, n_N) \\ & + \vec{\Gamma}_{j+1}(n_1, \dots, n_j + 1, \dots, n_N)]. \end{aligned} \quad (6)$$

Equations (1)–(6) can be numerically solved, subject to the normalization condition $\sum_{\mathbf{n}} \sigma(\mathbf{n}) = 1$.

^{a)}Electronic mail: uozumi@phys.aoyama.ac.jp

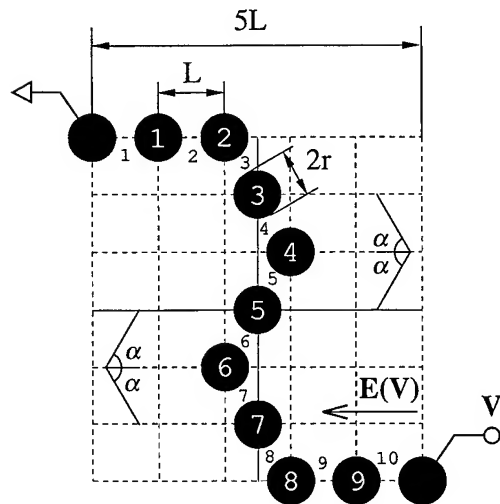


FIG. 1. Geometric design of the array for $\cos \alpha = 0.5$. The 9 islands and the 10 junctions are numbered in series from the left to right. The numbers of the backward junctions are 5 and 6 in the presence of external electric field $E(V)$.

The geometric design of the array proposed here is shown in Fig. 1. The array with a uniform junction resistance consists of nine spherical islands numbered 1–9 with radius r and spacing L (much larger than r). For the simplest calculations for the island potentials due to excess charges on

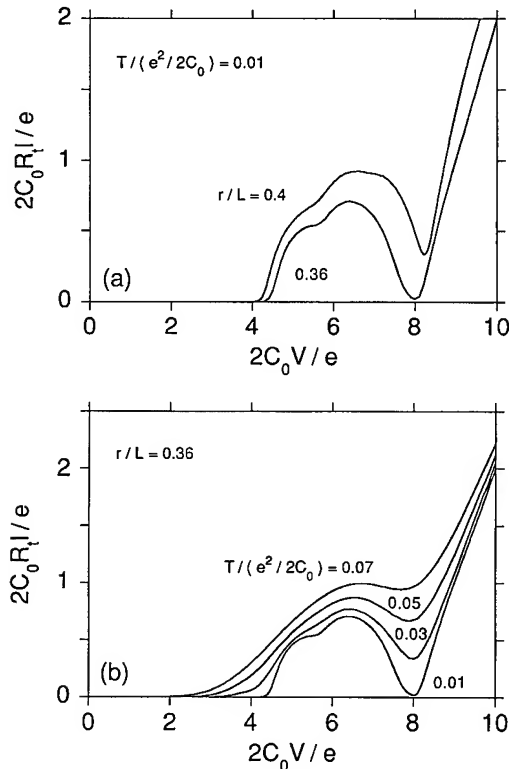


FIG. 2. Simulated I - V characteristics of the array. I - V curves for (a) two values of r/L , (b) several values of the temperature T . R_t and C_0 denote the sum of all the junction resistances and the self capacitance of island, respectively.

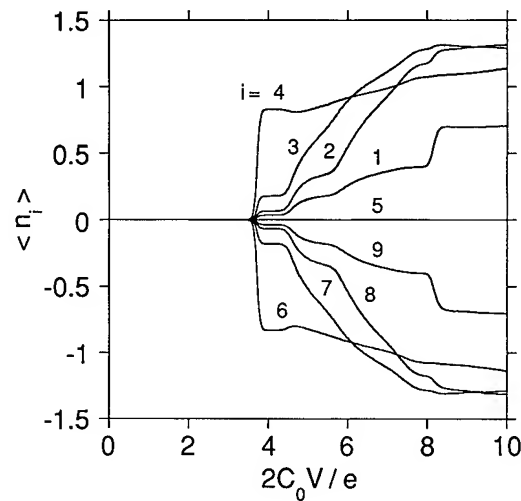


FIG. 3. Average number of excess electrons $\langle n_i \rangle$ on island i as a function of the external voltage V . The parameters are $r/L = 0.36$ and $T/(e^2/C_0) = 0.01$.

islands, the external electrodes, or leads, are also spherical with radius r , so that mutual capacitance of leads and islands (except islands 1 and 9) enters very little into our model. However, external field $E(V)$ is assumed to be altered in proportion to external voltage V , i.e., $E(V) = V/5L$. Tunneling is assumed to occur only between nearest-neighbor islands (numbered consecutively). For the potential difference between two electrodes related to the j th junction, $\Delta\phi_j = |\kappa_j V|$, due to the external field, we use the approximation $\Delta\phi_j = EL$ for $(j = 1, 2, 9, 10)$ and $\Delta\phi_j = EL \cos \alpha$ for $(j = 3, 4, \dots, 8)$. The internal potential of island i , U_i , is estimated by using the method of multiple images.

The results of the simulations by the Monte Carlo method are shown in Fig. 2. In Fig. 2(a), the system with larger r/L , i.e., in the case of more closely located islands, shows that the slope is smaller in the NDR characteristics. In Fig. 2(b), we show the temperature dependence of the I - V curve. An increase of the temperature washes out the negative slope.

In order to understand how the negative slope is produced, we discuss the electron-hole traffic flow in our nonlinear ten-junction array. In usual one-dimensional junction arrays with uniform junction resistance, at low voltages and temperatures, electron and hole traffic flow is restricted in each end of the junctions because of the Coulomb blockade. This effect often causes the so-called Coulomb gap in the I - V characteristics. Once a single electron and a single hole enter the array, they can flow from each side of the array to the other without any restriction until they meet in the middle of the array and annihilate together. On the other hand, in our nonlinear junction array with uniform junction resistance, the electron-hole traffic flow is restricted not only in edge junctions 1 and 10 but also in junctions 5 and 6 because junctions 5 and 6 go in the direction opposite to the external electric field, that is, negative transport voltage is applied across them. Since, as for an electron transfer, the

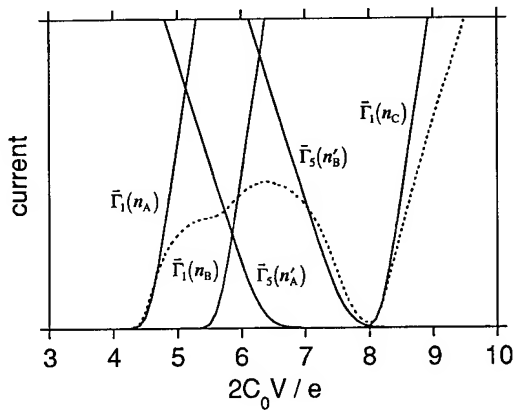


FIG. 4. Comparing the five electron hopping rates with the simulated I - V curve. The solid and dashed lines are the electron hopping rates and the I - V curve, respectively. The parameters are $r/L=0.36$ and $T/(e^2/C_0)=0.01$.

electron hopping rates across junctions 1 and 5 are much slower than any other possible electron hopping rates, only these rates chiefly contribute to the average transport time τ required for an electron to travel from the edge to the middle of the array. If, in a steady transport process, an electron enters the array under electron configuration \mathbf{n} and then another excess electron on island 4 hops across junction 5 under redistributed electron configuration \mathbf{n}' , then the average current can be written in the following form:

$$I \equiv \frac{1}{\tau} \sim \frac{1}{1/\tilde{\Gamma}_1(\mathbf{n}) + 1/\tilde{\Gamma}_5(\mathbf{n}')}. \quad (7)$$

This expression is good only for voltages small so that no $\langle n_i \rangle$ are much greater than one. In order to evaluate the average current from Eq. (7), one should choose the most probable \mathbf{n} and \mathbf{n}' according to the external voltages (refer to Fig. 3). We assume that electron configurations \mathbf{n} and \mathbf{n}' vary from \mathbf{n}_A and \mathbf{n}'_A to \mathbf{n}_C and \mathbf{n}'_C with increasing the external voltage as follows:

$$\begin{aligned} \text{set A: } \mathbf{n}_A &= \{0, 0, 0, 1, 0, -1, 0, 0, 0\} \quad \text{and} \\ \mathbf{n}'_A &= \{0, 0, 1, 1, 0, -1, -1, 0, 0\}, \\ \text{set B: } \mathbf{n}_B &= \{0, 0, 1, 1, 0, -1, -1, 0, 0\} \quad \text{and} \\ \mathbf{n}'_B &= \{0, 1, 1, 1, 0, -1, -1, -1, 0\}, \\ \text{set C: } \mathbf{n}_C &= \{0, 1, 1, 1, 0, -1, -1, -1, 0\} \quad \text{and} \\ \mathbf{n}'_C &= \{0, 1, 1, 2, 0, -2, -1, -1, 0\}. \end{aligned} \quad (8)$$

One may wonder why, first of all, a set of $\{0, 0, 0, 0, 0, 0, 0, 0, 0\}$ and $\{0, 0, 0, 1, 0, -1, 0, 0, 0\}$ is not chosen here. This is because

no transport process represented by this set happens since the negative transport voltage across junction 5 inhibits an excess electron on island 4 from hopping across junction 5 until a new second electron enters the array and then pushes it forward with Coulomb interaction. Set A represents just this transport process. In Fig. 4, the five rates $\tilde{\Gamma}_1(\mathbf{n}_A)$, $\tilde{\Gamma}_5(\mathbf{n}'_A)$, $\tilde{\Gamma}_1(\mathbf{n}_B)$, $\tilde{\Gamma}_5(\mathbf{n}'_B)$, and $\tilde{\Gamma}_1(\mathbf{n}_C)$ are plotted with the simulated I - V curve. One finds that our crude estimations of Eq. (7) partially fit the simulated I - V curve well. In a transport process represented by set B, with increasing the external voltage, $\tilde{\Gamma}_1(\mathbf{n}_B)$ increases but $\tilde{\Gamma}_5(\mathbf{n}'_B)$ decreases. According to Eq. (7), the current depends on the slower rate. Hence, at lower voltages the current is dominated by $\tilde{\Gamma}_1(\mathbf{n}_B)$ but at higher voltages the current is dominated by $\tilde{\Gamma}_5(\mathbf{n}'_B)$. When $\tilde{\Gamma}_1(\mathbf{n}_B)$ overcomes $\tilde{\Gamma}_5(\mathbf{n}'_B)$, the current starts to decrease. This negative slope lasts until $\tilde{\Gamma}_5(\mathbf{n}'_B)$ is overcome by the rate of another new electron that enters the array, i.e., $\tilde{\Gamma}_1(\mathbf{n}_C)$. That is why NDR appears. The negative slope can thus be seen at voltages where $\tilde{\Gamma}_1(\mathbf{n}_B) \leq \tilde{\Gamma}_5(\mathbf{n}'_B) \leq \tilde{\Gamma}_1(\mathbf{n}_C)$.

In conclusion, we have explored the possibilities that non-linear ten-junction arrays of nine ultrasmall islands containing two backward junctions show NDR characteristics. If the island radius r is taken to be 5 nm as it is in Fig. 2(b), the charging energy $e^2/2C_0 \sim 0.1$ eV and the junction resistances $R_j \sim 10$ G Ω (consistent with one-dimensional calculations for a barrier height of 1 eV), then one can see clear NDR at tens of kelvins.

ACKNOWLEDGMENT

This work was supported by the Grant-in-Aid for Scientific Research on Priority Area by the Ministry of Education, Science, Sports and Culture.

¹K. Mullen, E. Ben-Jacob, R. C. Jaklevic, and Z. Schuss, Phys. Rev. B **37**, 98 (1988).

²T. G. Miller and R. Reifenberger, Phys. Rev. B **50**, 3342 (1994).

³M. Amman, K. Mullen, and E. Ben-Jacob, J. Appl. Phys. **65**, 339 (1989).

⁴H. Nakashima and K. Uozumi, Jpn. J. Appl. Phys. **34**, L1659 (1995).

⁵K. K. Likharev, N. S. Bakhvalov, G. S. Kazach, and S. I. Serdyukova, IEEE Trans. Magn. **25**, 1436 (1989).

⁶U. Geigenmüller and G. Schön, Europhys. Lett. **10**, 765 (1989).

⁷Single Charge Tunneling Coulomb Blockade Phenomena in Nanostructures: Charge Tunneling Rates in Ultrasmall Junctions, edited by H. Grabert and M. H. Devoret (Plenum, New York, 1992), pp. 21 and 107.

Patterning of Langmuir–Blodgett film with ultrahigh vacuum-scanning tunneling microscope/atomic force microscope

H. Hamanaka, T. Ono, and M. Esashi

Faculty of Engineering, Tohoku University, Aza Aoba, Aramaki, Aoba-ku, Sendai 980-77, Japan

(Received 9 March 1997; accepted 10 March 1997)

We performed a patterning of a stearic acid Langmuir–Blodgett (LB) film on a Si substrate with ultrahigh vacuum scanning probe microscope. The patterning was carried out with a conducting Si cantilever probe by applying positive or negative bias voltages to the sample. When a negative bias (-10 V) was applied, one monolayer of the LB film was extracted. On the other hand, when a positive sample bias ($+10$ V) was applied, the scanned area was protruded. This extraction mechanism is explained in terms of the force caused by permanent and induced dipole moments. Further, SiO_2 was selectively deposited on the patterned area using SiCl_4 and H_2O . © 1997 American Vacuum Society. [S0734-211X(97)02404-9]

I. INTRODUCTION

Electronic devices have been scaled down by the development of microtechnology. In order to minimize the pattern size, many efforts, such as the exposure by short wavelength lights and the improvements of the resist, have been made. A pattern which is smaller than light wavelength, however, can not be delineated by photolithography in principle. Nanoscale lithography has been achieved using electron beam (EB), ion beam (IB), and x-ray lithography. Recently, a scanning probe microscope (SPM) lithography technique has been studied. SPM is superior to other lithography techniques as follows. It does not have the problem of proximity effect in EB exposure. Since it uses atomic level local interaction between the tip and the sample, it can even ultimately manipulate individual atoms. To achieve nanoscale patterning using these techniques, the thickness of the resist has to be very thin and pin-hole free. As a resist which satisfies these requirements, a hydrogen-terminated Si surface, inorganic resist, self-assembled monolayer film has attracted attention.

Hydrogen-terminated Si surface is chemically stable and has some resistivity against oxidation even in air. Some reports have been made on nanoscale patterning using this surface property. Lyding *et al.*¹ have made nanoscale oxidation patterns on a hydrogen-terminated Si surface using a ultrahigh-vacuum (UHV) scanning tunneling microscope (STM). After hydrogen is extracted from the surface, the area is exposed to oxygen and then achieved an atomic scale oxidation pattern. Sugimura *et al.*² demonstrated the SPM induced anodization. They have used a HF treated hydrogen-terminated Si surface and operated a SPM on this surface in air. This oxide pattern was used as a mask to etch Si in order to achieve a Si nanopattern. They also tried to make a pattern with electroless plating by noting the difference in chemical properties between the hydrogen-passivated surface and the oxidized surface.³ Ono *et al.*⁴ have made a selective deposition of Al with diMethyl–aluminum–hydride (DMAH) on the SPM modified hydrogen-terminated Si surface. It is considered that the selective growth of Al is due to the existence of free electrons on the hydrogen-terminated surface and the

chemical reaction between the surface terminated hydrogen and the methyl group of the DMAH.

Langheinrich *et al.*⁵ have performed nanoscale patterning by exposing AlF_3 -doped LiF film with EB. This inorganic film acts as a self-developing resist.

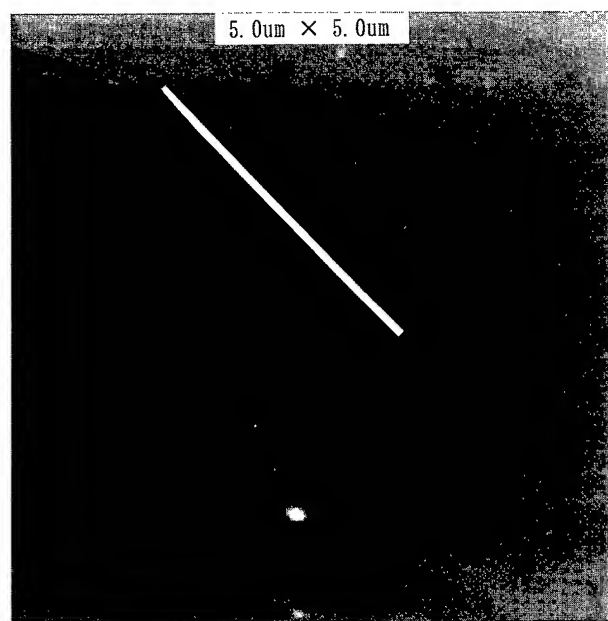
Self-assembled monolayer (SAM) films are formed on the surface with regularly arranged molecules, and very thin films (monolayer) can be achieved. A SAM can be patterned by exposing it to ultraviolet light or an electron beam. Using these differences of surface chemical properties, it is possible to form a pattern by a deposition technique. For these reasons, a SAM is considered to be a useful material for the lithography resist. Calvert *et al.*⁶ showed that it is possible to make a pattern with deep UV light in a organosilane SAM film, such as phenyl–tri–chloro–silane (PTCS) and subsequently they formed the Ni pattern by electroless plating.

Langmuir–Blodgett (LB) film is also a SAM film and its patterning has been reported. The thickness of the LB film can be controlled by a monolayer and its surface can be changed to both hydrophobic and hydrophilic. Barraud *et al.*⁷ have carried out a patterning of ω -tricosenoic acid LB film by polymerizing it with a EB. Stockman *et al.*⁸ made a pattern in ω -tricosenoic acid LB film by exposing it with a STM and made a Au pattern of about 15 nm in width.

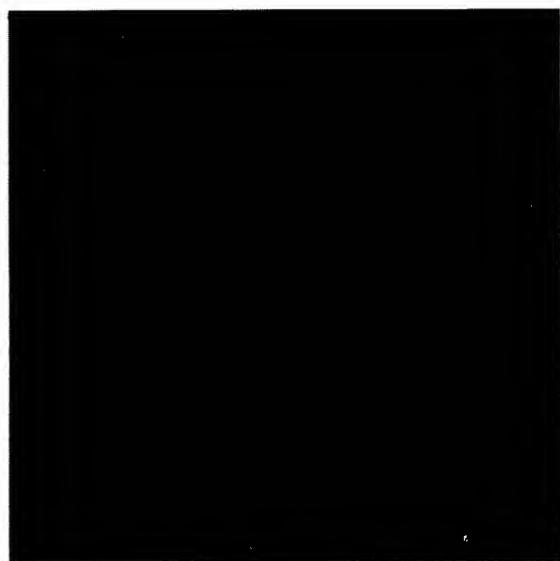
We performed the patterning of a stearic acid LB film by extracting molecules with an electric field using UHV SPM.

II. EXPERIMENT

Our experiments were made on a *n*-type Si(100), 0.02–0.03 Ω cm wafer. The sample was cleaned with the standard (RCA) method without a HF treatment. The surface is hydrophilic because of the natural oxide surface film. We use a stearic acid LB film in our experiments. Stearic acid is an amphiphilic molecule which has a normal chain expressed as follows: $\text{CH}_3(\text{CH}_2)_{16}\text{COOH}$. A chloroform solution of stearic acid was spread on an aqueous solution containing H_2SO_4 and NaHCO_3 , and the temperature of this solution was kept at 19 °C. The surface layer was compressed to form



(a)

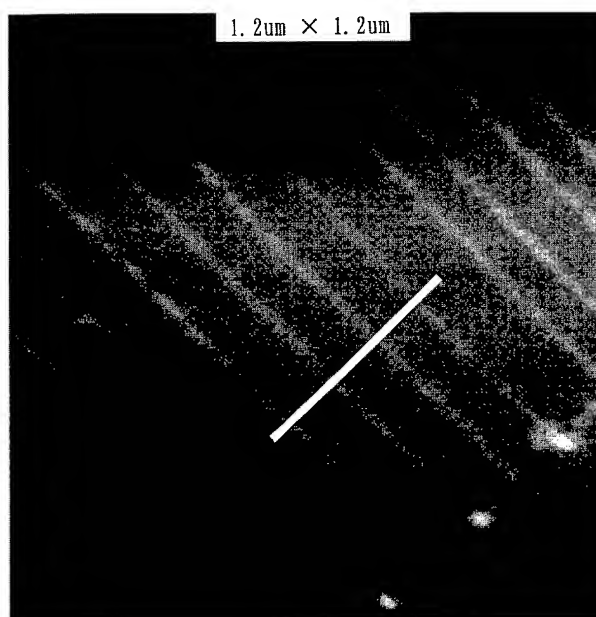


(b)

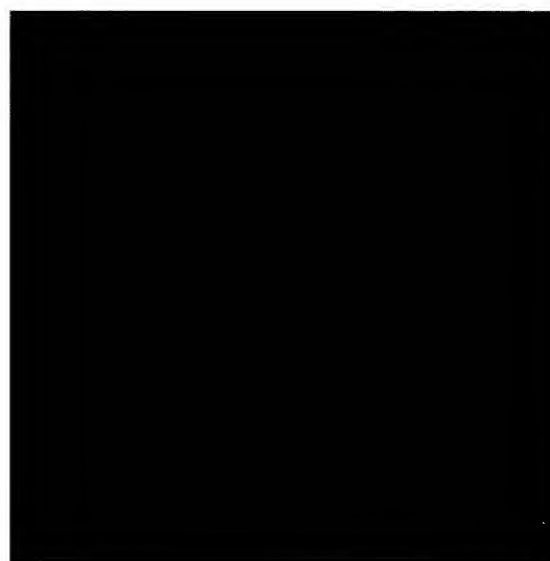
FIG. 1. (a) AFM image of LB film patterned with -10 V of sample bias with Si cantilever. Dark rectangular area shows removed area of one monolayer of LB film. (b) A cross sectional profile of the pattern shown in (a). The depth of the patterned area is about 1 nm.

a closely packed monolayer at 25 dyn cm^{-1} . On a hydrophobic surface, a LB film is not formed in the first down stroke. In the succeeding up stroke, a monolayer film is deposited on the surface with its hydrophilic group towards the sample. After this procedure, hydrophilic and hydrophobic groups are stacked alternately.

We used UHV-STM/atomic force microscope (AFM) (JEOL JSTM 4500 XT) for the patterning of the LB film. The patterning was performed under the base pressure of 1.3×10^{-8} Pa. We could apply a bias voltage to the sample for patterning the LB film using a Si cantilever probe.



(a)



(b)

FIG. 2. (a) AFM image of LB film patterned with $+10$ V of sample bias with Si cantilever. Bright line is the patterned region which seems to be protruded. (b) A cross section of the pattern shown in (a). The height of the pattern is about 1 nm.

III. RESULTS AND DISCUSSION

A. Patterning of LB film

First, one monolayer of LB film was deposited on a Si substrate. After the deposition of the LB film, the sample was introduced into the UHV-STM/AFM chamber. Two rectangular patterns and a line pattern were drawn using a Si cantilever with a sample bias voltage of -10 V, and a tunneling current of 10 nA. The scan directions are parallel to the line pattern. Figure 1(a) shows an AFM image of the patterned area, and this image was taken by the same canti-

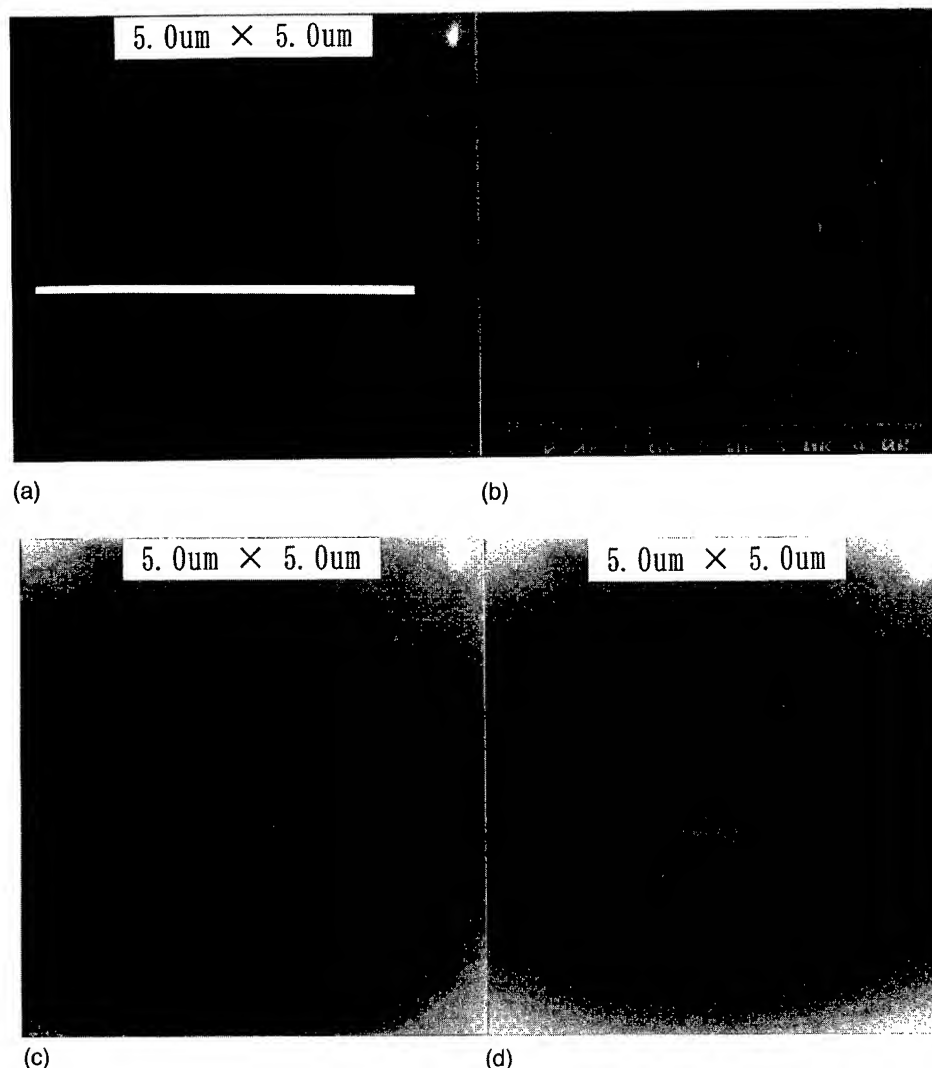


FIG. 3. AFM images of the surface after patterning five monolayers of LB film. (a) The area scanned with Si cantilever applying negative sample bias (-10 V). (b) A cross section of the pattern shown in (a). The depth is about 1 nm. (c) The smaller area than (a) scanned with positive sample bias ($+10$ V). (d) The larger area than (a) scanned with negative sample bias (-10 V).

lever, used for the patterning, and the patterned area was observed to be depressed. The depth measured by the AFM was about 1 nm as shown in Fig. 1(b). This depth was almost the same as the defect of the LB film, and if the slant of the LB film molecules is considered, it is thought that one monolayer of LB film was extracted, though the length of the stearic acid LB film is about 2 nm. Next, a patterning was performed with a positive bias voltage of $+10$ V and a tunneling current of 10 nA. Figure 2(a) shows the patterned area and it seemed to be protruded. The height of the pattern was about 1.2 nm by the AFM measurement as shown in Fig. 2(b). These results imply some effects of the electric fields between the tip and the sample. To investigate this patterning mechanism in detail, the same experiment was performed on multilayers of LB film.

Five monolayers of the stearic acid LB film were deposited on the substrate. The surface is hydrophobic. At first, a rectangular area was scanned with a sample bias of -10 V

and a tunneling current of 10 nA. Figure 3(a) shows an AFM image of the patterned area of the LB film, which seems to have been extracted. From the measurement with an AFM as shown in Fig. 3(b), the depth was about 1 nm which is thought to be equivalent to one monolayer. Next, the smaller area inside the rectangular area shown in Fig. 3(a) was scanned with a positive bias of $+10$ V. This region was observed to be a bright image as shown in Fig. 3(c). In this AFM image, the brighter area means a higher area and this implies the added LB film in the previously patterned area. Further, the larger area than that previously removed [Fig. 3(a)] was scanned with a negative sample bias voltage of -10 V. Then the protruded area [Fig. 3(c)] did not change and only the LB film newly patterned area was removed [Fig. 3(d)]. This newly patterned region was also equivalent to one monolayer from the measurement with the AFM. The area once patterned and the extracted area which came to an even number of layers did not show any changes. From these

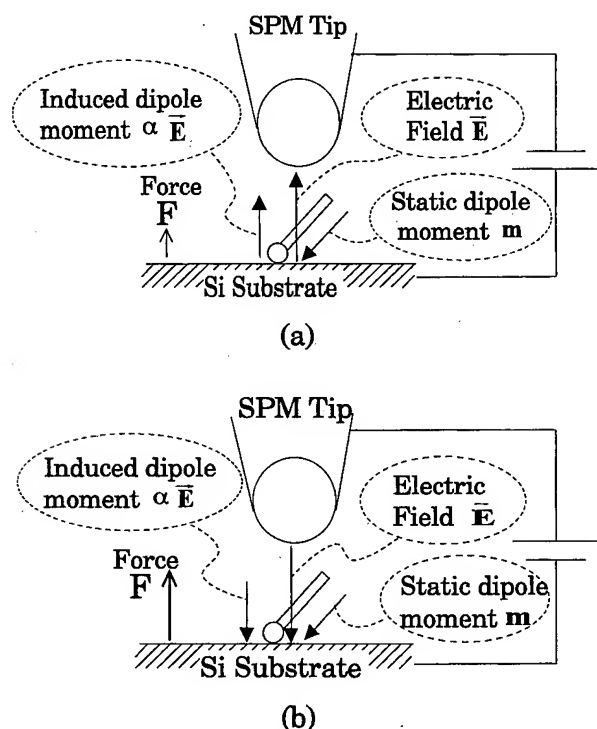


FIG. 4. Schematic illustration of extraction mechanism of LB film. (a) When sample bias is positive, the direction of induced dipole moment and static dipole moment are opposite. (b) When sample bias is negative, static and induced dipole moments are the same in their direction.

results, it can be said that only one monolayer of the LB film was removed, when the odd number of layers of the LB film were patterned with a negative sample bias.

To explain these phenomenon, we have framed a model as follows (Fig. 4). The LB film has polar molecules with

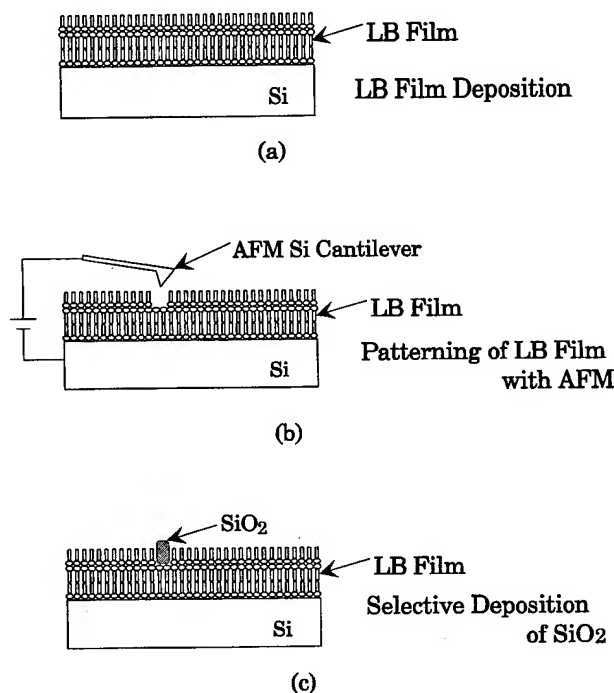


FIG. 5. Schematic procedure of selective deposition of SiO_2 using SiCl_4 and H_2O . (a) Odd number of monolayers of LB film is deposited on a Si substrate. (b) LB film is patterned with negative sample bias (-10 V) using Si cantilever. (c) SiO_2 grow selectively on the patterned area which is hydrophilic.

long hydrocarbon chains, which have permanent dipole moments. When an even number of layers are stacked, the permanent dipole moment is canceled. In the case of an odd number of layers, LB films have polarity. In this case, if the positive bias voltage is applied to the sample, the direction of

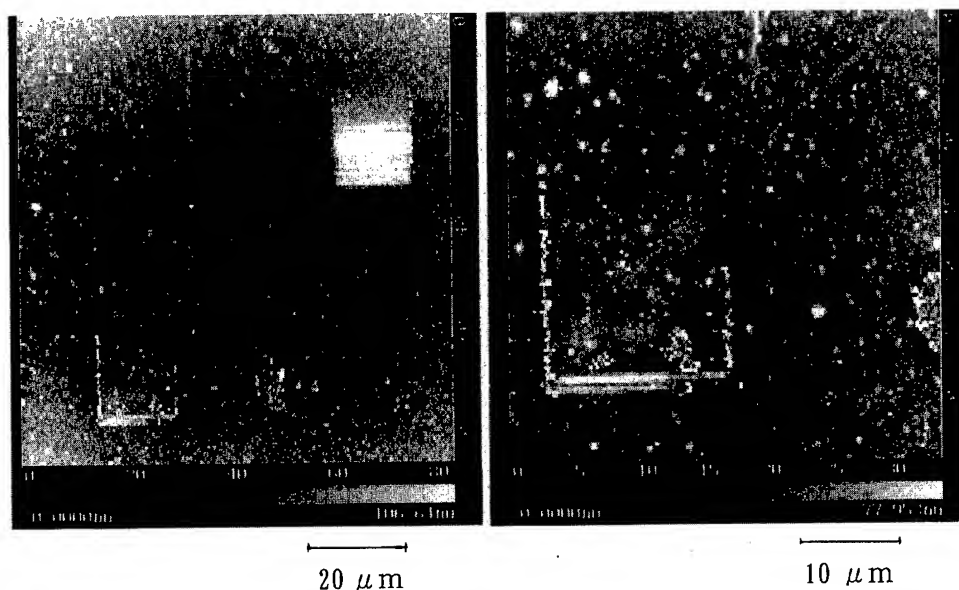


FIG. 6. (a) AFM image of SiO_2 pattern. Bright region (rectangular and line pattern) shows the SiO_2 pattern. Other region seems to be rough because LB film could not be deposited very well. (b) Magnification of the area shown in (a).

the induced dipole moment by an electric field is opposite to the permanent dipole moment and thus are canceled. On the other hand, when a negative bias voltage is applied, permanent and induced dipole moments have the same direction. This causes a large force to the molecule of the LB film and extracts one monolayer of the LB film.

B. Selective growth of SiO₂

With these results, the possibility of patterning the LB film has been recognized. Using the patterned LB film, selective growth of SiO₂ was performed. Figure 5 shows a principle to make the SiO₂ pattern. SiO₂ can be deposited using SiCl₄ and H₂O at room temperature.^{9,10} In this technique, surface adsorbed water plays an important role. On hydrophobic surfaces, SiO₂ film does not grow because there is no adsorbed water. Previously, SiO₂ selective growth was performed on the HF treated hydrogen-terminated Si surface patterned with a STM. The STM scanned area becomes hydrophobic because surface terminated hydrogen is removed and the area is oxidized. Following treatment of SiCl₄ and H₂O made selective SiO₂ deposition. Since the surface nature of the LB film (hydrophobic or hydrophilic) can be controlled, it was possible to form selective SiO₂ patterns.

Five monolayers of the LB film were deposited on the Si substrate, and this surface was patterned with a sample bias voltage of -10 V using a conducting AFM cantilever tip. As explained above, only one monolayer of the LB film on the patterned area was extracted. This area was considered to be hydrophilic as shown in Fig. 5(b). Then the sample was introduced into the vacuum chamber and the vapor deposition of SiO₂ was carried out at 1 Torr of SiCl₄ and 1 Torr of H₂O at room temperature in 2 min. SiO₂ grew selectively on the area where one monolayer of the LB film was extracted.

The bright area in Fig. 6 shows the deposition area of SiO₂. The other area was rough because the LB film could not be deposited smoothly. In order to achieve finer pattern, the LB film should be deposited more smoothly.

IV. CONCLUSION

Patterning of LB film with a UHV-SPM was carried out. The difference of sample bias polarity resulted in the protruded or depressed pattern. It is considered that the extraction phenomenon of the LB film depends on the force acting on the LB monolayer by the permanent and induced dipole moment. We have made a SiO₂ pattern using a different chemical nature of the patterned and the nonpatterned areas. SiO₂ grew on the SPM patterned hydrophilic area selectively.

ACKNOWLEDGMENTS

The authors would like to thank Professor T. Osa, Dr. J. Amzai, and Nishimura for their help in the LB film preparation.

¹J. W. Lyding, G. C. Abeln, and T. C. Shen, *J. Vac. Sci. Technol. B* **12**, 3735 (1994).

²H. Sugimura and N. Nakagiri, *Jpn. J. Appl. Phys.* **32**, 3406 (1995).

³H. Sugimura and N. Nakagiri, *Appl. Phys. Lett.* **66**, 1430 (1995).

⁴T. Ono, H. Hamanaka, and M. Esashi, *Thin Solid Films* (to be published).

⁵W. Langheinrich and H. Beneking, *Nanolithography: A Borderland Between STM, EB, IB, and X-Ray Lithographies* (Kluwer, Dordrecht, 1994), pp. 53-56.

⁶W. J. Dressick and J. M. Calvert, *Jpn. J. Appl. Phys.* **32**, 5829 (1993).

⁷A. Barraud, C. Rosilio, and A. Ruaudel-Teixier, *Solid State Technol.* **120** (1979).

⁸L. Stockman, G. Neuttiens, and C. Van Haensendonck, *Appl. Phys. Lett.* **62**, 2935 (1993).

⁹D. J. Ehrlich and J. Melngailis, *Appl. Phys. Lett.* **58**, 2675 (1991).

¹⁰N. Awaya and Y. Arita, *Jpn. J. Appl. Phys.* **27**, 1172 (1988).

Molecular organization of azobenzene derivatives at the liquid/graphite interface observed with scanning tunneling microscopy

P. C. M. Grim, P. Vanoppen, M. Rücker, S. De Feyter, S. Valiyaveetil,^{a)} G. Moessner,^{a)} K. Müllen,^{a)} and F. C. De Schryver^{b)}

Department of Chemistry, Laboratory of Molecular Dynamics and Spectroscopy, University of Leuven (KUL), Celestijnenlaan 200-F, 3001 Heverlee, Belgium

(Received 8 September 1996; accepted 15 February 1997)

Physisorbed monolayers of azobenzene derivatives were studied with a scanning tunneling microscope at the liquid/graphite interface. Three different compounds, namely, 4,4'-di(dodecyloxy) azobenzene ($C_{12}(AZO)C_{12}$), 5-[ω -(4'-dodecyloxy-4-azobenzene-oxy)-dodecyloxy] isophthalic acid ($C_{12}(AZO)C_{12}ISA$), and 4,4'-bis(ω -[3,5-bis(carboxylato)phenyl-1-oxy] dodecyloxy) azobenzene ($ISAC_{12}(AZO)C_{12}ISA$) have been investigated. In all cases monolayers could be observed with submolecular resolution at the liquid/graphite interface, allowing one to identify the azobenzene, as well as the other parts of the molecules. For each monolayer structure a molecular model could be composed with a good correspondence to the experimental data. Differences in the observed monolayer structures could be related to the chemical structure of the investigated compounds. The introduction of an isophthalic acid (ISA) headgroup has a profound influence on the monolayer structure because of its capability of hydrogen bond formation with other ISA groups or with solvent molecules. In addition, for the *trans* isomer of $C_{12}(AZO)C_{12}ISA$ solvent codeposition was observed. The compound $C_{12}(AZO)C_{12}ISA$ was used to investigate the effect of *cis-trans* isomerization on the molecular packing of the resulting monolayers. Under the influence of uv light, the azobenzene groups isomerize from the *trans* isomer to the thermodynamically less stable *cis* isomer. This isomerization process significantly changes the composition and structure of the observed monolayers. © 1997 American Vacuum Society. [S0734-211X(97)05604-7]

I. INTRODUCTION

Self-assembled monolayers have proven to be useful for surface modification, since they affect surface properties such as wetting, adhesion, and lubrication.¹ Chemisorbed self-assembled monolayers are spontaneously formed at the liquid/solid interface by chemisorption of molecules carrying a specific end group which can strongly bind to a certain solid surface. Well known examples are alkanethiols on gold¹⁻⁸ and silanes on silicon dioxide.^{1,9-11} These self-assembled monolayers are highly crystalline and the molecules are tilted with respect to the surface normal.

On the other hand, physisorbed self-assembled monolayers, in which the individual molecules are oriented parallel to the substrate surface, can be prepared as well. These monolayers are also spontaneously formed at the liquid/solid interface. In the absence of a specific interaction between the molecules and the solid surface, the monomolecular adlayer is built up by physisorption of individual molecules. Eventually, they can form a two dimensional molecular lattice on the solid surface. If the molecules are densely packed, this pattern of oriented molecules can be reproducibly imaged with submolecular resolution using scanning tunneling microscopy (STM). A large variety of compounds has already been studied in this way, including alkanes, alcohols, fatty

acids, and many others.¹²⁻²⁸ Also, codeposition of solvent molecules has been recently observed.²⁹

So far, the structural investigation of azobenzene derivatives with scanning probe techniques has been limited to chemisorbed self-assembled monolayers³⁰⁻³² and Langmuir-Blodgett films.³³

In general, molecules and polymers containing azobenzene groups have received attention due to their interesting photoresponsive behavior. The *cis-trans* isomerization effect of azobenzene chromophores offers a possible approach for, e.g., optical data storage.³⁴ On the other hand, some azobenzene derivatives exhibit liquid crystalline behaviour,^{35,36} suggesting long-range ordering of these compounds.

We have studied the monolayer formation behaviour of three azobenzene derivatives, namely, 4,4'-di(dodecyloxy) azobenzene [$C_{12}(AZO)C_{12}$, see Fig. 1(a)], 5-[ω -(4'-dodecyloxy-4-azobenzene-oxy)-dodecyloxy] isophthalic acid [$C_{12}(AZO)C_{12}ISA$, see Fig. 1(b)] and 4,4'-bis(ω -[3,5-bis(carboxylato)phenyl-1-oxy] dodecyloxy) azobenzene [$ISAC_{12}(AZO)C_{12}ISA$, see Fig. 1(c)]. In addition, structural changes within the monolayer upon *trans-cis* isomerization have been investigated for $C_{12}(AZO)C_{12}ISA$.

II. EXPERIMENT

The synthesis of $C_{12}(AZO)C_{12}ISA$ has been recently reported.²⁹ The synthesis of the compounds $C_{12}(AZO)C_{12}$ and $ISAC_{12}(AZO)C_{12}ISA$ is described below in steps (a) through (d). Steps (a), (b), and (d) were conducted under a

^{a)}Also at: Max-Planck Institute for Polymer Research, P.O. Box 3148, D-55021 Mainz, Germany.

^{b)}Author to whom correspondence should be addressed.

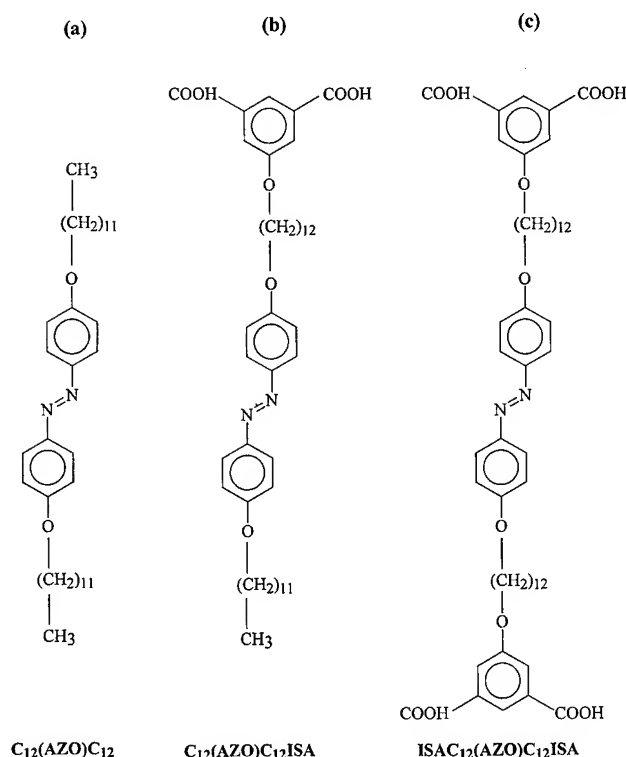


FIG. 1. Chemical structures of the investigated compounds: (a) 4,4'-di(dodecyloxy) azobenzene [$C_{12}(AZO)C_{12}$], (b) 5-[ω-(4'-dodecyloxy-4-azobenzene-oxy)-dodecyloxy] isophthalic acid [$C_{12}(AZO)C_{12}ISA$], and (c) 4,4'-bis(ω-(3,5-bis(carboxylato)phenyl-1-oxy) dodecyloxy) azobenzene [$ISAC_{12}(AZO)C_{12}ISA$].

dry Ar atmosphere. Unless mentioned otherwise, 1H and ^{13}C nuclear magnetic resonance (NMR) spectra were recorded on a Varian Gemini 200 spectrometer at room temperature. Chemical shifts are reported in ppm and are referenced to residual proton containing solvent [$\delta(CDCl_3) = 7.26$; $\delta(DMSO - d_6) = 2.49$]. Mass spectra were determined on a ZAB2-SE-FPD instrument. All commercially available reagents were used without further purification.

Synthesis of (a) dimethyl 5-(ω-bromo dodecyloxy)-isophthalate: 1,12-dibromo dodecane (8.0 g, 0.024 mol), powdered K_2CO_3 (3.37 g, 0.024 mol), a catalytic amount of KI, and 100 ml abs. N,N-dimethyl formamide were placed in a two-necked round bottom flask equipped with a reflux condenser, a pressure equalizing funnel, and a magnetic stirrer. The mixture was heated up to 80 °C, then dimethyl 5-hydroxy isophthalate (2.56 g, 0.012 mol) in 30 ml of abs. N,N-dimethyl formamide was added dropwise over a period of 6 h. After stirring the solution for another 10 h at 80 °C, the solvent was removed. The crude product was taken up in dichloromethane and washed twice with water and a saturated solution of sodium chloride. After drying over magnesium sulfate and removal of dichloromethane, the product was purified by column chromatography on silica gel (eluent: ethyl acetate/*n*-heptane=1:5). The pure product is a white crystalline solid; yield: 3.1 g (56%). An alternative procedure, frequently applied in the cases of shorter chain dibromo alkanes,³⁷ using a sevenfold excess of the dibromo

alkane which is distilled off after the reaction, afforded the product in lower yields of around 45%. These lower yields are probably due to some degree of decomposition during the distillation at temperatures of over 160 °C (5×10^{-2} bar).

1H -NMR($CDCl_3$): δ 8.26 (s, 1 H), 7.73 (s, 2 H), 4.05 (t, 2 H, $J = 3$ Hz), 3.95 (s, 6 H), 3.41 (t, 2 H, $J = 4$ Hz), 1.94–1.72 (m, 4 H), 1.62–1.23 (m, 16 H); ^{13}C -NMR($CDCl_3$): δ 166.6, 159.7, 132.2 (2), 123.2 (2), 120.3 (2), 76.9, 69.1, 52.8, 34.4, 33.3, 30.0, 29.9, 29.8, 29.6, 29.2, 28.6, 26.4; EI-MS m/z 458.1 [$(M + 1)^+$].

Synthesis of (b) 4,4'-bis(ω-(3,5-bis(methylcarboxylato)phenyl-1-oxy) dodecyloxy) azobenzene: 4,4'-dihydroxy azobenzene (0.71 g, 3.3 mmol) and the ω-bromo diester from step (a) (3.03 g, 6.6 mmol) were mixed with powdered K_2CO_3 (2.29 g, 0.016 mol) and a catalytic amount of KI and the mixture was suspended in 50 ml of abs. N,N-dimethyl formamide. The suspension was stirred for 8 h at 80 °C under a dry Ar atmosphere. The workup was carried out as described in step (a). The crude product was purified by double recrystallization from methanol. Yield: 2.74 g (85.5%).

1H -NMR($CDCl_3$): δ 8.25 (s, 2 H), 7.86 (d, 4 H, $J = 4$ Hz), 7.73 (s, 4 H), 6.97 (d, 4 H, $J = 4$ Hz), 4.06 (t, 8 H, $J = 2$ Hz), 3.93 (s, 12 H), 1.90–1.67 (m, 8 H), 1.60–1.15 (m, 32 H); ^{13}C -NMR($CDCl_3$): δ 166.7 (4), 161.6 (2), 159.7 (2), 147.4 (2), 132.2 (4), 124.7 (4), 122.6 (2), 120.1 (4), 115.1 (4), 68.2 (2), 67.9 (2), 52.7 (4), 30.0 (8), 29.7 (4), 29.5 (4), 25.2 (2), 25.0 (2); EI-MS m/z 967.4 [$(M + 1)^+$].

(c) 4,4'-bis(ω-(3,5-bis(carboxylato)phenyl-1-oxy) dodecyloxy) azobenzene [$ISAC_{12}(AZO)C_{12}ISA$; Fig 1(c)]: The methyl ester from step (b) (2.5 g, 2.6 mmol) was hydrolyzed by refluxing it with 3.5 equivalents of potassium hydroxide in a mixture of ethanol and water (2:1) for 8 h. After evaporating the alcohol, the alkaline aqueous solution was rendered neutral by addition of semiconcentrated hydrochloric acid up to a pH of 4–5. The product precipitated as a yellow solid, which was taken up in diethyl ether. The organic phase was washed twice with water and dried over magnesium sulfate. The crude acid was purified by double recrystallization from methanol. Yield: 1.72 g (73%); 1H -NMR($DMSO - d_6$): δ 8.10 (s, 1 H), 7.81 (d, 4 H, $J = 4$ Hz), 7.65 (s, 2 H), 7.07 (d, 4 H, $J = 4$ Hz), 4.08 (t, 8 H, $J = 2$ Hz), 1.95–1.65 (m, 8 H), 1.55–1.15 (m, 32 H); ^{13}C -NMR ($DMSO - d_6$, 300 MHz, 60 °C): δ 166.1 (4), 160.7 (2), 158.6 (2), 146.0 (2), 132.5 (4), 123.7 (4), 121.9 (2), 118.9 (4), 114.7 (4), 68.0 (2), 67.8 (2), 28.6 (8), 28.4 (4), 28.2 (4), 25.2 (2), 25.1 (2); FD-MS m/z 911.2 [$(M + 1)^+$].

(d) 4,4'-di(dodecyloxy) azobenzene [$C_{12}(AZO)C_{12}$; Fig. 1(a)]: 4,4'-dihydroxy azobenzene (5 g, 23.36 mmol) and 1-bromododecane (11.63 g, 46.72 mmol) were mixed with powdered K_2CO_3 (16.14 g, 0.117 mol) and a catalytic amount of KI and the mixture was suspended in 300 ml of abs. N,N-dimethyl formamide. The suspension was stirred for 8 h at 80 °C. The workup was carried out as described in step (a). Purification was achieved by double recrystallization of the crude product from methanol. Yield: 10.41 g (81%); 1H -NMR($CDCl_3$): δ 7.86 (d, 4 H, $J = 4$ Hz), 6.99

(*d*, 4 H, *J* = 4 Hz), 4.05 (*t*, 4 H, *J* = 3 Hz), 1.92–1.74 (*m*, 4 H), 1.57–1.09 (*m*, 36 H), 0.90 (*t*, 6 H, *J* = 2.5 Hz); ^{13}C -NMR(CDCl₃): δ 161.7 (2), 147.5 (2), 124.8 (4), 115.2 (4), 66.9 (2), 32.4 (2), 30.0 (6), 29.8 (6), 29.7 (2), 26.5 (2), 23.1 (2), 14.6 (2); FD-MS *m/z* 550.8 (*M*⁺).

The STM data were obtained using a Discoverer instrument (Topometrix Inc., Santa Barbara, CA), operating under ambient conditions. An external pulse/function generator (Hewlett Packard HP 8111 A) was used to supply the bias voltage and to occasionally apply short voltage pulses to the STM tip. Samples were prepared as follows. First the compounds were dissolved in either 1-phenyloctane (Aldrich, 99%) or 1-undecanol (Aldrich, 99%), at concentrations of about 1 mg/ml. Then a drop of solution was applied to the basal plane of a piece of highly oriented pyrolytic graphite (HOPG, grade ZYB, Advanced Ceramics Inc., Cleveland, OH). STM tips were electrochemically etched from platinum/iridium wire (80%/20%; diameter 0.2 mm), in a 2 N KOH/6 N NaCN solution in water. The tip was immersed into the liquid drop applied to the basal plane of the graphite and approached the interface. Imaging was performed in the current imaging mode, with bias voltages of 0.5 V up to 1.5 V (sample negative) and an average tunneling current of 1 nA. STM images of the graphite surface, obtained at low tunneling voltages, were used for calibration. The data acquisition speed, limited by the STM instrument, was fairly low. For a 200×200 image frame, approximately 7 s were needed. For the *cis-trans* isomerization experiments a CAMAG universal lamp (366 nm) was used to illuminate the solution on the graphite surface.

III. RESULTS AND DISCUSSION

First, the observed monolayers of the three different compounds presented in Fig. 1 will be reported. The solution applied to the HOPG surface was stored in the dark for at least several days to ensure a high *trans* isomer content, and consequently, to facilitate the observation of monomolecular adlayers of *trans* isomer molecules.

Figures 2(a) and 2(b) show STM images of the adlayer structure of C₁₂(AZO)C₁₂ molecules deposited from 1-phenyloctane on the graphite surface.

The two dimensional molecular lattice of the physisorbed monolayer can be clearly observed. In Fig. 2(a) even the direction of the alkyl chains is visible. The brighter regions in the STM image originate from the azobenzene moieties and correspond to a higher tunneling current. From the literature,^{12,20,21} it is well known that for molecules containing both alkyl and aryl moieties, the tunneling current above aromatic groups is increased with respect to the alkyl parts of the molecules. This has been observed for a large variety of organic adsorbates on graphite. On the basis of the STM results a molecular model for the two dimensional packing of the C₁₂(AZO)C₁₂ molecules on the graphite surface is proposed, along with a unit cell of the structure. Figure 2(c) displays the molecular model that was composed on the basis of these STM observations. All molecules are oriented in the same direction and are interdigitated over the full length of

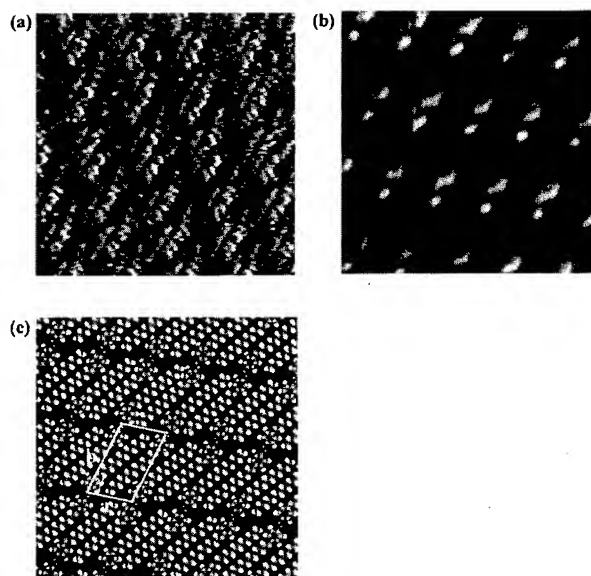


FIG. 2. (a) STM image of a monolayer of C₁₂(AZO)C₁₂ molecules adsorbed from a solution in 1-phenyloctane. The direction of the alkyl chains can be clearly recognized in the image, as well as the positions of the azobenzene groups. The image size equals 8.6 × 8.6 nm². A line defect is visible in the horizontal direction in the middle of the image. (b) STM image of a monolayer of C₁₂(AZO)C₁₂ molecules adsorbed from a solution in 1-phenyloctane. The contrast of the aromatic groups is predominant in the image. The size of the image equals 6.2 × 6.2 nm². (c) Molecular model built on the basis of the STM image presented in (b). The unit cell, as indicated in the model, contains one C₁₂(AZO)C₁₂ molecule. The unit cell parameters *a*, *b*, and *γ*, as defined in the molecular model, are 1.5 ± 0.1 nm, 2.0 ± 0.2 nm, and 67 ± 3°, respectively.

the alkyl chains. The unit cell parameters *a*, *b* and *γ*, as defined in the molecular model, have been determined to be 1.5 ± 0.1 nm, 2.0 ± 0.2 nm, and 67 ± 3°, respectively. Every benzene ring in the molecular model corresponds to a bright spot in the STM image.

The observed monolayer structure of C₁₂(AZO)C₁₂ISA [Fig. 1(b)], turned out to be remarkably different. The introduction of the isophthalic acid (ISA) head group allows the formation of hydrogen bonds between neighboring isophthalic acid head groups on the graphite surface. Therefore, the packing of the monolayer is to a large extent determined by the relative position of the ISA head groups. Unfunctionalized isophthalic acid derivatives have already been studied by STM.³⁸ X-ray data from 3D supramolecular self-assembled structures are available as well.^{39–41}

Figure 3(a) shows the monolayer structure of the *trans*-C₁₂(AZO)C₁₂ISA, after a drop of solution of the compound in 1-undecanol was applied to the graphite surface. As can be seen from the figure, a lamellar monolayer structure is formed. In order to indicate the observed pattern, arrows have been inserted underneath the STM image. The isophthalic acid head groups of the molecules are situated along the direction of the arrows marked ISA. Like the AZO groups, they appear as bright spots in the STM image. The spacing between the ISA head groups is 0.97 nm. The spacing between the two marked ISA rows is 1.9 nm, measured perpendicularly to the direction of the lamellae. This distance

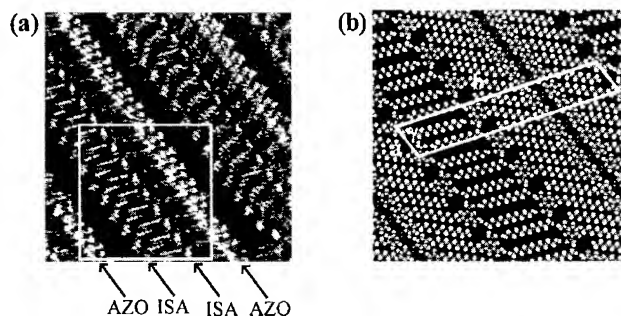


FIG. 3. (a) STM image of a monolayer of $C_{12}(AZO)C_{12}ISA$ molecules adsorbed from a solution in 1-undecanol. A lamellar monolayer structure is formed which is characterized by the rows of ISA head groups and AZO chromophores marked by arrows below the image. Codeposition of 1-undecanol, used as a solvent for $C_{12}(AZO)C_{12}ISA$, is visible in the image. The solvent molecules are located between the rows of ISA head groups. The size of the image equals $13.0 \times 13.0 \text{ nm}^2$. (b) Molecular model of the indicated area in the STM image in (a), along with a proposed unit cell. The $C_{12}(AZO)C_{12}ISA$ molecules are interdigitated, leading to a closely packed monolayer structure. The unit cell contains two 1-undecanol molecules and two *trans*- $C_{12}(AZO)C_{12}ISA$ molecules. The unit cell parameters a , b , and γ , as defined in the model, are $0.97 \pm 0.10 \text{ nm}$, $6.6 \pm 0.4 \text{ nm}$, and $66 \pm 4^\circ$, respectively.

is much too large for an effective hydrogen bonding to occur between the ISA head groups of two molecules, situated in adjacent lamellae. A closer inspection of the image contrast in between the two ISA rows shows that this space is occupied by smaller molecules. Based on the measured distance between the ISA rows and the angle relative to the direction of the lamellae at which these molecules appear, it can be concluded that these molecules are actually solvent molecules, coadsorbed with the *trans*- $C_{12}(AZO)C_{12}ISA$ molecules. The azobenzene moieties are located in the somewhat broader bright bands, indicated with the arrows marked AZO. A molecular model for the monolayer structure along with a unit cell is proposed in Fig. 3(b). Again, there is a good correspondence between the model and the molecular adlayer represented within the white square at the bottom of the STM image in Fig. 3(a). The unit cell contains two $C_{12}(AZO)C_{12}ISA$ and two 1-undecanol molecules. The unit cell parameters a , b , and γ were assessed to be $0.97 \pm 0.10 \text{ nm}$, $6.6 \pm 0.4 \text{ nm}$, and $66 \pm 4^\circ$, respectively. The $C_{12}(AZO)C_{12}ISA$ molecules are fully interdigitated, as can be seen from the molecular model. They form extended lamellae in the direction of the arrows, which are separated by lamellae consisting of 1-undecanol molecules. Occasionally, it was possible to observe pairs of $C_{12}(AZO)C_{12}ISA$ lamellae in between which no solvent molecules were incorporated. This fact will be illustrated together with the discussion of the *cis-trans* isomerization results (*vide infra*).

Also monolayers formed by the compound $ISAC_{12}(AZO)C_{12}ISA$ [Fig. 1(c)], have been observed at the liquid/graphite interface. Both ends of the molecules are now terminated with ISA groups, which are capable of hydrogen bond formation with either other ISA groups or solvent molecules. In Fig. 4(a) a STM image of the observed monolayer structure is presented. Due to the presence of the second ISA group, the possibility of interdigitation, as in the case of

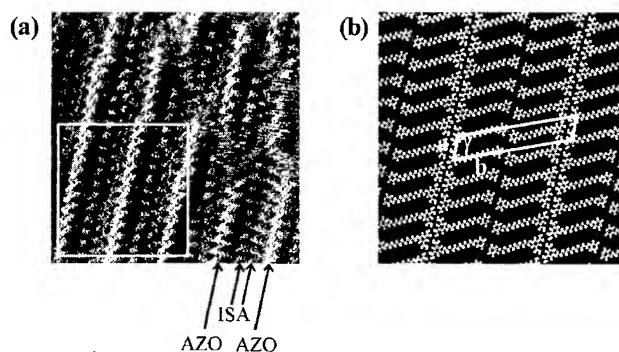


FIG. 4. (a) STM image of a monolayer of $ISAC_{12}(AZO)C_{12}ISA$ molecules adsorbed from a solution in 1-undecanol. Again, a lamellar monolayer structure is formed which is characterized by the rows of ISA head groups and AZO chromophores marked by arrows below the image. Some domains, which are shifted with respect to one another, are visible. The overall image contrast is relatively poor, probably due to a reduced monolayer stability at the liquid/graphite interface compared to the $C_{12}(AZO)C_{12}$ and $C_{12}(AZO)C_{12}ISA$ monolayers. The image size equals $15.8 \times 15.8 \text{ nm}^2$. (b) Molecular model of the indicated area in the STM image in (a), along with a proposed unit cell. The model shows noninterdigitated $ISAC_{12}(AZO)C_{12}ISA$ molecules. The unit cell contains one $ISAC_{12}(AZO)C_{12}ISA$ molecule. The unit cell parameters a , b , and γ , as defined in the molecular model, are $1.0 \pm 0.2 \text{ nm}$, $5.1 \pm 0.9 \text{ nm}$, and $70 \pm 5^\circ$, respectively.

$C_{12}(AZO)C_{12}ISA$ and $C_{12}(AZO)C_{12}$, is now excluded. However, a lamellar structure is still observed. The rows of ISA head groups and AZO groups are indicated below the STM image by arrows. A remarkable observation is the complete absence of 1-undecanol solvent molecules in the respective monolayer. There are indications that the presence of solvent molecules in the monolayer depends on the concentration of the solution. A molecular model from the area indicated in the STM image, including a unit cell, is presented in Fig. 4(b). The unit cell parameters a , b , and γ equal $1.0 \pm 0.2 \text{ nm}$, $5.1 \pm 0.9 \text{ nm}$, and $70 \pm 5^\circ$, respectively. Free space between the $ISAC_{12}(AZO)C_{12}ISA$ molecules within one lamella is clearly visible in the model. This fact might account for the reduced spatial resolution in the STM image, as compared to the results obtained with $C_{12}(AZO)C_{12}$ and $C_{12}(AZO)C_{12}ISA$, because the molecules might be less immobilized than in a densely packed monolayer. Whether the free space is occupied by solvent molecules or whether the bare graphite surface is exposed, cannot be concluded at present from the experimental data.

In addition, we report on the light-induced *trans-cis* isomerization of the compound $C_{12}(AZO)C_{12}ISA$ and the resulting monolayer structure. Upon irradiation of the solution on the graphite surface with light at a wavelength of 366 nm, the azobenzene chromophore changes its geometry (Fig. 5) and a photostationary mixture of *trans* and *cis* isomers is formed. When shielded from light, the *cis* isomer spontaneously converts back to the thermodynamically more stable *trans* isomer. Since we are able to simultaneously image both isomers, the dark reaction from *cis* to *trans* in the adlayer is slower than the time needed for the experiment. More detailed information regarding the light-induced *cis-*

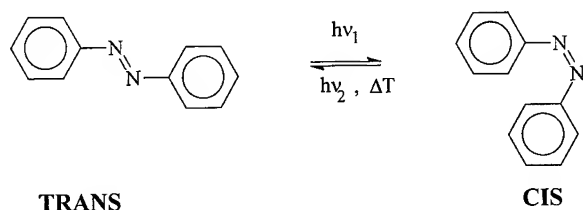


FIG. 5. Schematic drawing illustrating the *cis-trans* isomerization of an azobenzene chromophore.

trans isomerization of $C_{12}(AZO)C_{12}ISA$ on graphite is reported elsewhere.²⁹

Figure 6(a) shows a STM image of the observed monolayer structure after irradiating a drop of the $C_{12}(AZO)C_{12}ISA$ solution in 1-undecanol on the graphite surface. Three domains can be identified in the image. Domain 1 corresponds to the monolayer structure of the *trans* isomer of $C_{12}(AZO)C_{12}ISA$, already discussed previously [Figs. 3(a) and 3(b)]. Domains 2 and 3, however, exhibit a different monolayer structure, not observed before. This monolayer structure can be explained by a model of the pure *cis* isomer in which no solvent molecules are incorporated. Figure 6(b) displays the molecular model from the indicated area in domain 3. The model is composed by *cis*- $C_{12}(AZO)C_{12}ISA$ molecules which are not interdigitated, as opposed to the situation in the monolayers of *trans*- $C_{12}(AZO)C_{12}ISA$. This different packing constitutes an easy way of discerning *cis* and *trans* domains on the graphite surface. The fact that no solvent molecules are incorporated in the monolayer structure is probably due to a different spatial orientation of the ISA head groups, which makes H bond formation with molecules in adjacent lamellae more favorable than H bond formation with solvent molecules.

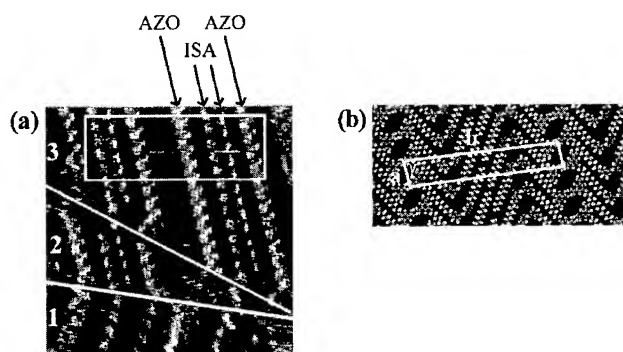


FIG. 6. STM image of an ordered monolayer structure consisting of *cis* and *trans* $C_{12}(AZO)C_{12}ISA$ domains. Rows of ISA head groups and AZO chromophores are indicated above the image. Domain 2 and 3 consist of *cis*- $C_{12}(AZO)C_{12}ISA$ molecules, whereas domain 1 contains *trans*- $C_{12}(AZO)C_{12}ISA$ molecules. The image size equals $13.4 \times 13.4 \text{ nm}^2$. (b) Molecular model of the indicated area in the STM image in (a), along with a proposed unit cell for the *cis* monolayer structure. The *cis* monolayer structure does not incorporate solvent molecules as in the case of the *trans* monolayer [Figs. 3(a) and 3(b)]. The unit cell dimensions a , b , and γ have been determined as $1.03 \pm 0.10 \text{ nm}$, $6.0 \pm 0.4 \text{ nm}$, and $85 \pm 4^\circ$, respectively. The unit cell contains two *cis*- $C_{12}(AZO)C_{12}ISA$ molecules.

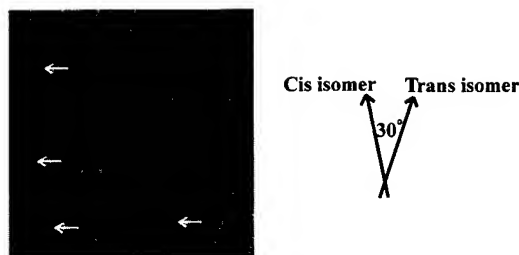


FIG. 7. STM image showing coexisting *cis* and *trans* domains of $C_{12}(AZO)C_{12}ISA$ molecules (size: $40 \times 40 \text{ nm}^2$). The *cis* and *trans* lamellae are oriented in the direction of the respective arrows next to the image. The angle between *cis* and *trans* lamellae is always 30° . This is due to the alignment of the alkyl chains of the *cis* and *trans* isomers in their respective lamellae.

Another way to distinguish the *cis* and *trans* $C_{12}(AZO)C_{12}ISA$ domains is based on the direction of their respective lamellae. Figure 7 is a STM image of an area of $40 \text{ nm} \times 40 \text{ nm}$, showing an assembly of *cis* and *trans* $C_{12}(AZO)C_{12}ISA$ domains. All the *cis* domains are rotated by the same angle with respect to the *trans* domains. A closer look at the models of the *cis* and *trans* $C_{12}(AZO)C_{12}ISA$ isomers leads to the conclusion that in the *cis* domains the alkyl chains between the ISA head groups and the AZO chromophore are lying in the same direction as the alkyl chains in the *trans* domains. This fact accounts for the fixed angle (30°) between the *cis* and *trans* domains. The STM image in Fig. 7 further illustrates that solvent incorporation inbetween *trans* lamellae is not a prerequisite for monolayer formation. Occasionally, areas are found where solvent incorporation is absent. This is indicated in the STM image by the white arrows. Furthermore, the *cis* domains have never been observed without the presence of *trans* domains on the graphite surface. This is probably due to a combination of factors, like the relative abundancy of the two species in solution and the monolayer stability of both isomers at the liquid/graphite interface.

IV. CONCLUSION

We have shown that azobenzene derivatives can be successfully imaged with STM at the liquid/graphite interface with submolecular resolution. Individual azobenzene chromophores and isophthalic acid groups can be easily identified. Structural changes between the observed monolayers could be related to the chemical structure and nature of the investigated compounds. The molecules with ISA head groups [$C_{12}(AZO)C_{12}ISA$ and $ISAC_{12}(AZO)C_{12}ISA$] both adopt a lamellar configuration at the liquid/graphite interface. The ISA head groups force the molecules into a lamellar configuration enabling hydrogen bonding between ISA head groups and between ISA head groups and solvent molecules. This leads to a closely packed, stable monolayer structure. However, in the case of $ISAC_{12}(AZO)C_{12}ISA$, a reduced monolayer stability and image contrast was observed. This can probably be attributed to the unoccupied space between the molecules within one lamella, as a result of which the molecular mobility is increased. Codeposition

of solvent molecules was observed in the case of *trans*-C₁₂(AZO)C₁₂ISA. When solvents capable of hydrogen bonding are employed, the phenomenon of solvent codeposition might be used to control the structure of the resulting monolayer. Moreover, it might be an effective way for immobilizing small molecules, which are not able to form stable monolayers by themselves, due to their high mobility.

In addition, we have shown that *cis-trans* isomerization has a profound influence on the resulting monolayer structure in the case of C₁₂(AZO)C₁₂ISA. Monolayers of reagent and product of a reversible photoinduced reaction have been simultaneously imaged and identified at the liquid/graphite interface.

ACKNOWLEDGMENTS

The authors thank the "Fonds Kollektief Fundamenteel Onderzoek" (FKFO) and the "Federale Diensten voor Wetenschappelijke, Technische en Culturele Aangelegenheden" (DWTC) for continuing financial support through the "InterUniversitaire AttractiePool" (IUAP, Grant No. IV-11). S. De Feyter is a predoctoral fellow of the "Fonds voor Wetenschappelijk Onderzoek" (FWO), and P. Vanoppen thanks the "Vlaams Instituut ter bevordering van het Technisch-Wetenschappelijk Onderzoek in de Industrie" (IWT), for a predoctoral fellowship.

¹For an overview see: A. Ulman, *An Introduction to Ultrathin Films* (Academic, San Diego, CA, 1991).

²C.D. Bain, E. B. Troughton, Y.T. Tao, J. Evall, G. M. Whitesides, and R. G. Nuzzo, *J. Am. Chem. Soc.* **111**, 321 (1989).

³P. E. Laibinis, M. A. Fox, J. P. Folkers, and G. M. Whitesides, *Langmuir* **7**, 3167 (1991).

⁴L. Bertilsson and B. Liedberg, *Langmuir* **9**, 141 (1993).

⁵K. L. Prime and G. M. Whitesides, *Science* **252**, 1164 (1991).

⁶C. A. Alves, E. L. Smith, and M. D. Porter, *J. Am. Chem. Soc.* **114**, 1222 (1992).

⁷G. E. Poirier and E. D. Pylant, *Science* **272**, 1145 (1996).

⁸E. Delamarche and B. Michel, *Thin Solid Films* **273**, 54 (1996).

⁹L. Netzer and J. Sagiv, *J. Am. Chem. Soc.* **105**, 674 (1983).

¹⁰N. Tillman, A. Ulman, J. S. Schildkraut, and T. L. Penner, *J. Am. Chem. Soc.* **110**, 6136 (1988).

¹¹H. O. Finklea, L. R. Robinson, A. Blackburn, B. Richter, D. Allara, and T. Bright, *Langmuir* **2**, 239 (1986).

¹²Recently, a review paper appeared, concerning interfacial adsorbate imaging, imaging mechanisms and dynamical processes at the liquid/solid interface: D. M. Cyr, B. Venkataraman, and G. W. Flynn, *Chem. Mater.* **8**, 1600 (1996).

¹³J. P. Rabe and S. Buchholz *Phys. Rev. Lett.* **66**, 2096 (1991).

¹⁴J. P. Rabe and S. Buchholz, *Science* **253**, 424 (1991).

¹⁵A. Wawkuszewski, H.-J. Cantow, and S. N. Magonov, *Langmuir* **9**, 2778 (1993).

¹⁶N. Elbel, W. Roth, E. Günther, and H. von Seggern, *Surf. Sci.* **303**, 424 (1994).

¹⁷J.-C. Poulin, *Microsc. Microanal. Microstruct.* **5**, 351 (1994).

¹⁸M. Hibino, A. Sumi, and I. Hatta, *Jpn. J. Appl. Phys.* **1** **34**, 3354 (1995).

¹⁹M. Hibino, A. Sumi, and I. Hatta, *Thin Solid Films* **273**, 272 (1996).

²⁰S. Cincotti and J. P. Rabe, *Supramolecular Science* **1**, 7 (1994).

²¹K. Walzer and M. Hietschold, *J. Vac. Sci. Technol. B* **14**, 1461 (1996).

²²D. L. Patrick, V. J. Cee, and T. P. Beebe, Jr., *J. Phys. Chem.* **100**, 8478 (1996).

²³D. M. Walba, F. Stevens, D. C. Parks, N. A. Clark, and M. D. Wand, *Science* **267**, 1144 (1995).

²⁴B. Venkataraman, G. W. Flynn, J. L. Wilbur, J. P. Folkers, and G. M. Whitesides, *J. Phys. Chem.* **99**, 8684 (1995).

²⁵F. Stevens, D. J. Dyer, and D. M. Walba, *Angew. Chem. Int. Ed. Engl.* **35**, 900 (1996).

²⁶A. Stabel, R. Heinz, F. C. De Schryver, and J. P. Rabe, *J. Phys. Chem.* **99**, 505 (1995).

²⁷A. Stabel, R. Heinz, J. P. Rabe, G. Wegner, F. C. De Schryver, D. Corens, W. Dehaen, and C. Süling, *J. Phys. Chem.* **99**, 8690 (1995).

²⁸R. Heinz, A. Stabel, J. P. Rabe, G. Wegner, F. C. De Schryver, D. Corens, W. Dehaen, and C. Süling, *Angew. Chem.* **106**, 2154 (1994).

²⁹P. Vanoppen, P. C. M. Grim, M. Rücker, S. De Feyter, G. Moessner, S. Valiyaveetil, K. Müllen, and F.C. De Schryver, *J. Phys. Chem.* **100**, 19636 (1996).

³⁰H. Wolf, H. Ringsdorf, E. Delamarche, T. Takami, H. Kang, B. Michel, Ch. Gerber, M. Jaschke, H.-J. Butt, and E. Bamberg, *J. Phys. Chem.* **99**, 7102 (1995).

³¹W. B. Caldwell, D. J. Campbell, K. Chen, B. R. Herr, C. A. Mirkin, A. Malik, M. K. Durbin, P. Dutta, and K. G. Huang, *J. Am. Chem. Soc.* **117**, 6071 (1995).

³²M. Jaschke, H. Schönherr, H. Wolf, H.-J. Butt, E. Bamberg, M. K. Be-socke, and H. Ringsdorf, *J. Phys. Chem.* **100**, 2290 (1996).

³³R. Wang, L. Jiang, T. Iyoda, D. A. Tryk, K. Hashimoto, and A. Fujishima, *Langmuir* **12**, 2052 (1996).

³⁴T. Ikeda and O. Tsutsumi, *Science* **268**, 1873 (1995).

³⁵H. Akiyama, M. Momose, K. Ichimura, and S. Yamamura, *Macromolecules* **28**, 288 (1995).

³⁶Y. Tabe and H. Yokoyama, *Langmuir* **11**, 4609 (1995).

³⁷See, e.g., J. V. Crivello, M. Deptolla, and H. Ringsdorf, *Liq. Cryst.* **3**, 235 (1988).

³⁸K. Eichhorst-Gerner, A. Stabel, D. Declercq, J. P. Rabe, G. Moessner, S. Valiyaveetil, V. Enkelmann, and K. Müllen, *Angew. Chem. Int. Ed. Engl.* **35**, 1492 (1996).

³⁹M. Pfaadt, G. Moessner, D. Pressner, S. Valiyaveetil, C. Boeffel, K. Müllen, and H. W. Spiess, *J. Mater. Chem.* **5**, 2265 (1995).

⁴⁰S. Valiyaveetil, V. Enkelmann, and K. Müllen, *J. Chem. Soc. Chem. Commun.* **1994**, 2097 (1994).

⁴¹S. Valiyaveetil, V. Enkelmann, and K. Müllen, *Am. Chem. Soc., Polym. Div. Polym. Prepr.* **1995**, 552 (1995).

Formation of nanocolumn self-assembly by solvent polarity control

Xiangdong Chai, Wensheng Yang, Yunwei Cao, Yueshun Jiang, Ran Lu, and Tiejin Li
Department of Chemistry, Jilin University, Changchun 130023, People's Republic of China

Huixin He, Yongqiang Wang, and Zhongfan Liu
Center for Intelligent Materials, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, People's Republic of China

(Received 8 September 1996; accepted 7 May 1997)

The morphology of self-assembly formed by amphiphilic complementary components, 5-(4-dodecyloxybenzylidene)-2,4,6-(1H,3H)-pyrimidinetrione and 4-amino-2,6-didodecylamino-1,3,5-triazine, was changed from square and rectangle sheets into nanocolumns and helices corresponding to a recognition pattern change from a linear tape to a cyclic hexamer or a helically grown tape by varying the polarity of the solvent. © 1997 American Vacuum Society.
[S0734-211X(97)10404-8]

I. INTRODUCTION

Self-assembly, a process widely used in forming complex biological systems, is being extensively explored as an effective way to construct nanostructures for various applications.¹⁻³ Molecular recognition directed self-assembly via hydrogen bondings has received more attention than those of other weak or nondirectional interactions, such as Van de Waals, π -aromatic stacking, hydrophobic interactions, etc., due to its stronger force and directional nature.^{4,5} In multiple hydrogen bond associated self-assemblies positive cooperation often speeds formation and eliminates disarrangement of molecular components in self-assemblies. The self-assembly of complementary compounds, barbituric acid or cyanuric acid derivatives with melamine or pyrimidine derivatives, is such a system and it has been extensively studied by several groups.⁶⁻¹⁰ Up until now at least four molecular recognition patterns, linear tape, crinkled tape, cyclic hexamer,¹¹ and helically grown cyclic tape,¹² have been suggested. The basis of these pattern formation has been reported as the stereo demand of the substituents, other weak interactions, and unknown factors. In some cases it is hard to determine the recognition structure with experimental methods. As a result, it is difficult to predesign the stacking of molecular components and predict their macroscopic expression, especially when combined with properties and functions.

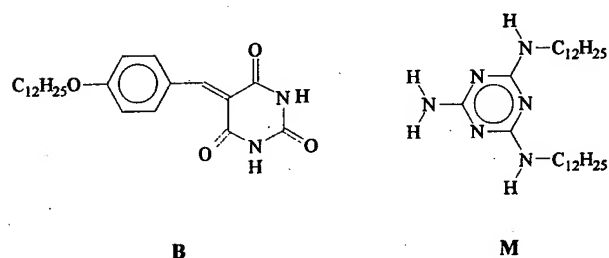
With the aim of learning more about the factors that influence the recognition patterns of the above mentioned complementary pairs and of obtaining control over their assembly morphologies, we studied the effect of polar solvent treatment on the self-assembly formed by complementary components, 5-(4-dodecyloxybenzylidene)-2,4,6-(1H,3H)-pyrimidinetrione (**B**) and 4-amino-2,6-didodecylamino-1,3,5-triazine (**M**). In another paper we reported that **B** and **M** recognize each other in the linear tape pattern in anhydrous chloroform, and the resulting self-assembly, when it was dropped onto a hydrophilic surface, exhibits morphologies of square and rectangular sheets with layered structures.¹³ Just like the components used in similar work,^{6,8,9} **B** and **M** are also amphiphilic compounds that we

have successfully prepared into alternating Langmuir-Blodgett films.^{14,15} we can imagine that their self-assembly should be sensitive to the surrounding polarity. By transmission electron microscopy (TEM) and atomic force microscope (AFM) observations, and ultraviolet (UV)-visible absorption and fluorescence measurements the self-assembly in a linear tape was turned into a cyclic hexamer pattern, and its morphology in square and rectangular sheets into nanofilaments and even helices when it was treated with an ethanol/water mixture, anhydrous ethanol, or methanol. This indicates that the polarity of the solvent is another important factor by which to determine the recognition patterns and the morphology of self-assembly. This finding may open the way in controlling the properties and functions combined with molecular components at the molecular level and yield new kinds of nanomaterials.

II. EXPERIMENT

The syntheses of **B** and **M** were reported in Ref. 16 and their molecular structures are shown in Fig. 1. Anhydrous chloroform was prepared by refluxing it overnight with P_2O_5 under argon. Absolute ethanol and methanol were obtained by the treating commercial absolute solvents (produced at Tianjin Chemicals) with magnesium and then distilling them afterward under dry argon.

The self-assembly was obtained by dissolving equivalent molar of **B** and **M** in anhydrous chloroform, refluxing them for 2 h under argon atmosphere, and then keeping them in the dark at room temperature (20 °C). The solubility of **M** is much higher than that of **B**. For general purposes we prepare the solution in the concentration of 5×10^{-4} M. For higher concentrations the undissolved **B** formed at the beginning of mixing will dissolve during refluxing with **M**, which is an indication of this self-assembling process. The resulting self-assembly exists as a suspension or colloid at room temperature. An ethanol or methanol treatment was conducted by adding it into the self-assembly suspension in dry chloroform or directly preparing the self-assembly in chloroform containing methanol or ethanol in the same way as in the dry

FIG. 1. Molecular structure of **B** and **M**.

chloroform. This type of treatment can be also realized by immersing the self-assembly on a substrate into methanol or ethanol.

Morphology observations were carried out with a JEOL TEM-1200EX and a Digital Instruments Nano-III AFM. Fluorescence spectra were measured on a Shimadzu RF-5000 spectrofluorophotometer. UV-visible absorption was recorded on a Shimadzu 365.

III. RESULTS AND DISCUSSIONS

Morphology observations with TEM of the sample dropped on a copper grid covered with an amorphous carbon film and stained with a saturated solution of uranyl acetate in ethanol/water revealed that the self-assembly, which is shown in Fig. 2, exists in the form of coils with an average length of about 5 μm and a diameter of 300 nm. These coils can be found in almost every zone within the area in Fig. 2. With SEM, AFM, and a polarity optical microscope, the self-assembly on the Si(111), quartz, glass, and highly oriented pyrolytic graphite (HOPG) surfaces, the self-assembly gives morphologies (seen in Fig. 3) of square and rectangular sheets several tens of micrometers long and several micrometers thick. The square and rectangular sheets are well characterized as a layer self-assembly of **B** and **M**, recognized in a linear tape pattern.¹³ We noticed that the only distinction between the sample preparations is the stain treat-



FIG. 2. TEM micrograph of a coil self-assembly of **B** and **M** on an amorphous carbon film. The sample was stained with a saturated solution of uranyl acetate in ethanol/water (50/50).



FIG. 3. SEM micrograph of a rectangular sheet self-assembly of **B** and **M** prepared by directly depositing the self-assembly suspension in CHCl_3 onto quartz.

ment for the TEM sample, to which an extra polar solvent treatment and heavy metal deposition were performed.

In order to examine staining effect, UV-absorption spectra of the self-assembly in a $\text{CHCl}_3/\text{EtOH}/\text{H}_2\text{O}$ (80/10/10) mixture solvent and anhydrous $\text{CHCl}_3/\text{EtOH}$, and even anhydrous EtOH, MeOH were recorded, and it was found that **B** decomposes gradually in water containing solvents, but in other anhydrous solvents, **B** does not decompose at all. Fur-

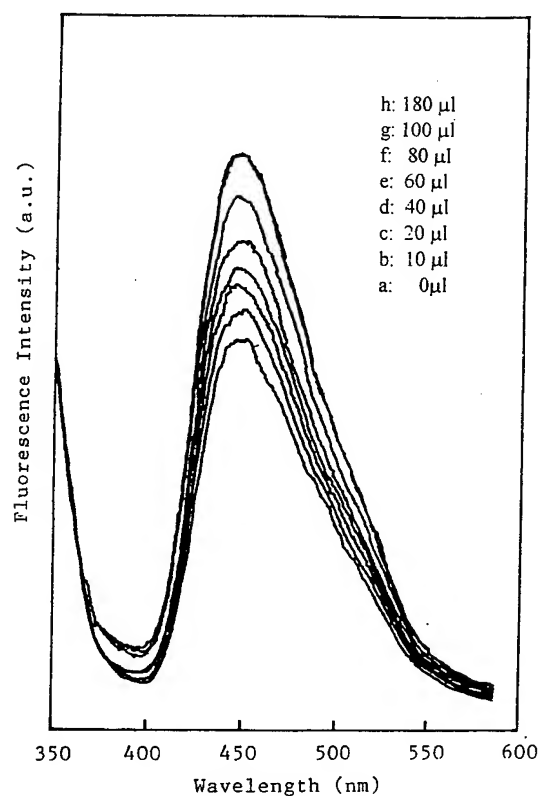


FIG. 4. Fluorescence spectra obtained by adding 0 (a), 10 (b), 20 (c), 40 (d), 60 (e), 80 (f), 100 (g), 180 (h) μl of anhydrous methanol into 5 ml of 5×10^{-5} M of the self-assembly in anhydrous chloroform. The emission is the ICT band excited at 306 nm.

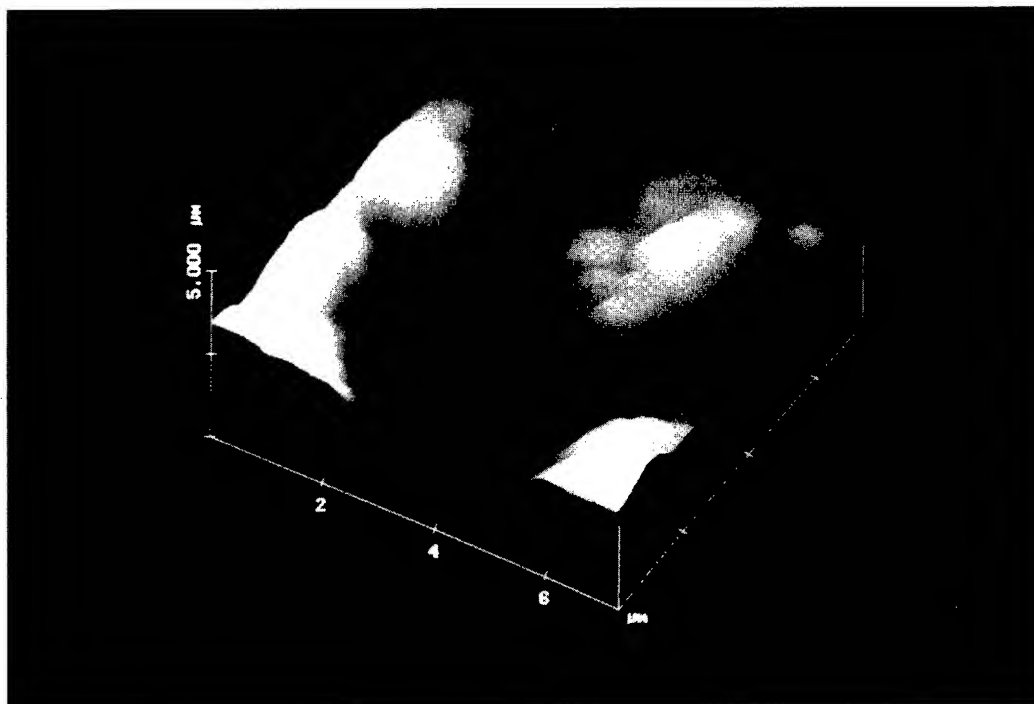


FIG. 5. An AFM image showing that the amorphous self-assembly of **B** and **M**, when destroyed by the AFM tip, is composed of a nanocolumn aggregation that is almost vertical to the substrate surface.

thermore, TEM observations on the samples stained with saturated anhydrous ethanol and methanol solutions of uranyl acetate proved that coil formation had no relation to water caused decomposition. The tendency for coil formation in the solvents is from ethanol/water (50/50) > methanol > ethanol, which is the exact polarity sequence of these three protonic solvents. This means that coil formation does depend on the polarity of the solvent. It is the polar solvents, methanol or ethanol, that change the self-assembly with sheet morphology into a coil structure; the stronger the polarity of the solvent, the stronger the tendency to coil. Heavy metal ion deposition was also confirmed as not causing coil formation since the coils can be also found through AFM observation on the self-assembly treated with anhydrous ethanol on the HOPG surface.

On careful observation of the coils we find the single strands that compose the coils are columns and not layered sheets; so coil formation from square or rectangular sheets corresponds to a transition of molecular stacking in the self-assembly. One possible way for the transition may be through the self-coiling of the linear recognition tape of **B** and **M**, which can be considered to be the basic unit of the layered sheet. In fact some helical assemblies of chiral compounds are formed in this way.¹⁷ If it is the case, this column should be a hollow tube with a relatively large diameter because the linear tape is rather rigid. However, the tube structure was not proven by AFM observations which yielded solid column images. So we can conclude that there must be a recognition pattern change for this self-assembly during the ethanol or methanol treatment. Since both the layered sheet and nanocolumn structures are of great importance in nano-

materials, it is important to have a understanding to the chemical and physical events in this transition.

We have proved that the fluorescence of component **B** is sensitive to the this self-assembling process,^{18,19} so we attempted to make the fluorescence measurements follow the processes by adding a small amount of anhydrous methanol to the **BM** self-assembly in anhydrous chloroform. Figure 4 shows the fluorescence spectra that recorded this process. The emission is the intramolecular charge transfer (ICT) band excited at 306 nm. It can be seen that the intensity increases with the addition of methanol, but with more than 100 μ l of methanol it does not increase further, which means that just a small amount of methanol is needed to induce the change of recognition pattern and additional methanol has no further effect on it.

To examine the structural variation caused by such a small amount of methanol, the resulting self-assembly suspension was dropped onto the HOPG surface for AFM observation. In this case, the common morphology of this sample, although the size is almost the same as the sheet formed in dry chloroform, is not in as regular a shape as that of the sheet and the top surface is not as flat as on the sheet. Moreover, it displayed rather amorphous features compared to the microcrystalline behavior of the rectangular sheet (it displays double refraction under a polar microscope). By destroying the center of this self-assembly with the AFM tip, a hole was made. Figure 5 is an AFM image around the hole. It can be seen that many columns with diameters in the nanometer scale range stand almost vertical to the substrate surface. Thus a conclusion can be made that this amorphous self-assembly is composed of the aggregation of columns

and it is due to the small amounts of methanol that dramatic changes in its pattern and morphology can be brought about.

The diameters of the columns in Fig. 5 vary from around 100 to 200 nm, and there is no definite periodic structure. In fact, sometimes TEM and AFM observations of the coils revealed that a single strand is composed of a much thinner filamentous structure, such as that shown at the branching position of the coil in Fig. 2. Therefore, we can conclude that the diameters of both the strands of the coils and the columns comprising the amorphous self-assemblies do not truly represent the size of the self-assembly in certain recognition patterns, but information on the recognition pattern is revealed by the columnar morphology in either the coils or the column aggregation. Accordingly, only the cyclic hexamer or the helically grown cyclic tape recognition pattern, which has the smallest round shaped unit, can be responsible for this shape. The nanocolumn should be composed of face-to-face stacking of the cyclic hexamers through π -aromatic interactions along the column long axis direction and the lateral aggregation through van de Waals and hydrophobic forces, or by the lateral aggregation of helically grown cyclic tapes. The coil morphology may be formed only when treated with much more methanol or ethanol, in which nanocolumns have a tendency to connect to each other in order to avoid the entropy decrease of the polar solvent.^{20,21} In fact, one can envision that in coil structures two strands have the largest overlap than in any other configuration. Further experiments to determine which of the two recognition patterns is responsible for the construction of column morphology are being performed.

IV. CONCLUSION

Polar solvents, ethanol and methanol, induce a change on the recognition pattern of the self-assembly of amphiphilic complementary components, 5-(4-dodecyloxybenzylidene)-2,4,6-(1H,3H)-pyrimidinetrione (**B**) and 4-amino-2,6-didodecylamino-1,3,5-triazine (**M**), from a linear tape to a cyclic hexamer or a helically grown cyclic tape. As a result, the morphology changed from square and rectangular sheets into nanocolumn aggregations and even coils. Previous research has shown that stereo demands are the main factor determining the recognition pattern of hydrogen bonded

complementary components;¹¹ here we have proved that the solvent could be another factor that can affect it. This may provide a way to manipulate molecular components and to modify their properties and functions in the macroscopic range.

ACKNOWLEDGMENT

The authors thank Professor Jean-Marie Lehn for initiating part of the syntheses of the compounds used here and for many helpful discussions.

¹J.-M. Lehn, *Angew. Chem. Int. Ed. Engl.* **29**, 1304 (1990).

²G. M. Whitesides, J. P. Mathias, and C. T. Seto, *Science* **254**, 1312 (1991).

³J. Lindsey, *New. J. Chem.* **15**, 153 (1991).

⁴J. G. Steven, C. Vicent, E. Fan, and A. D. Hamilton, *Angew. Chem. Int. Ed. Engl.* **32**, 119 (1993).

⁵M. C. Etter, *J. Phys. Chem.* **95**, 4601 (1991).

⁶J.-M. Lehn, M. Mascal, A. DeCian, and J. Fischer, *J. Chem. Soc. Perkin Trans. 2*, 461 (1992).

⁷S. K. Chang and A. D. Hamilton, *J. Am. Chem. Soc.* **110**, 1318 (1988).

⁸N. Kimizuka, T. Kawasaki, and T. Kunitake, *J. Am. Chem. Soc.* **115**, 4387 (1993).

⁹C. T. Seto and G. M. Whitesides, *J. Am. Chem. Soc.* **112**, 6409 (1990).

¹⁰W. S. Yang, S. G. Chen, X. D. Chai, Y. W. Cao, R. Lu, W. P. Chai, Y. S. Jiang, X. Y. Tang, T. J. Li, and J.-M. Lehn, *Synth. Met.* **71**, 2107 (1995).

¹¹J. A. Zerkowski, J. C. MacDonald, C. T. Seto, D. A. Wierda, and G. M. Whitesides, *J. Am. Chem. Soc.* **116**, 2382 (1994).

¹²N. Kimizuka, S. Fujikawa, J. Kuwakara, T. Kunitake, A. Marsh, and J.-M. Lehn, *J. Chem. Soc. Chem. Commun.* 2103 (1995).

¹³Y. W. Cao, X. D. Chai, H. X. He, J. Y. Sun, W. S. Yang, Y. S. Jiang, Z. F. Liu, and T. J. Li (unpublished).

¹⁴Y. Z. Ren, S. G. Chen, X. D. Chai, Y. W. Cao, W. S. Yang, R. Lu, G. Li, A. D. Lu, M. X. Gong, Y. B. Bai, T. J. Li, and J.-M. Lehn, *Synth. Met.* **71**, 1709 (1995).

¹⁵Y. Z. Ren, B. Zhao, X. D. Chai, and T. J. Li, *Thin Solid Films* **293**, 170 (1997).

¹⁶Y. W. Cao, X. D. Chai, S. G. Chen, Y. S. Jiang, W. S. Yang, R. Lu, M. Blanchard-Desce, T. J. Li, and J.-M. Lehn, *Synth. Met.* **71**, 1733 (1995).

¹⁷J.-H. Fuhrhop, S. Svenson, C. Boettcher, E. Rössler, and H.-M. Vieth, *J. Am. Chem. Soc.* **112**, 4307 (1990).

¹⁸R. Lu, S. G. Chen, X. D. Chai, Y. W. Cao, W. S. Yang, Y. S. Jiang, and T. J. Li, *Synth. Met.* **71**, 2035 (1995).

¹⁹W. S. Yang, X. D. Chai, R. Lu, Y. W. Cao, Y. S. Jiang, X. Y. Tang, T. J. Li, W. W. Wang, L. S. Fu, and H. J. Zhang, *Chem. Res. Chin. Univ.* **12**, 296 (1996).

²⁰K. A. Dill, *Science* **250**, 297 (1990).

²¹K. A. Sharp, A. Nicholls, R. F. Fine, and B. Honig, *Science* **252**, 106 (1991).

Writing and reading bit arrays for information storage using conductance change of a Langmuir-Blodgett film induced by scanning tunneling microscopy

K. Takimoto, R. Kuroda, S. Shido, S. Yasuda, H. Matsuda, K. Eguchi, and T. Nakagiri
Canon Research Center, Canon Inc., 5-1, Morinosato-Wakamiya, Atsugi, Kanagawa 243-01, Japan

(Received 12 September 1996; accepted 30 January 1997)

We demonstrate writing and reading bit arrays for information storage using an atomic force microscope with an electrically conducting probe and a medium consisting of a polyimide Langmuir-Blodgett film. In this system, data bits of 10 nm in diameter were written by application of voltage pulses with the conducting probe, and were read out by detecting the change in current. Topographical change was hardly observed at the data bits. A bit array of about 10^3 bits which contained information such as binary encoded characters was written and was successfully read out. The way to write and read the data bits with higher rate is discussed. © 1997 American Vacuum Society. [S0734-211X(97)11104-0]

I. INTRODUCTION

Application of scanning probe microscopy (SPM) to an information storage is promising, because the size of a data bit will be from submicron to subnanometer and is much smaller than that in any present storage device.¹ Hence it is expected that high density and huge capacity information storage can be achieved with a SPM. In order to realize the complete system, however, it is important to read and write the bit array with the rate as high as that of present storage devices.

In the SPM, a surface image is obtained by detecting various interactions between the tip of the probe and the surface of a sample when they are placed in close proximity. The gap between the tip and the sample surface has to be controlled precisely to keep it constant during scanning, in most cases. Therefore, it is predicted that high speed operation is difficult. The exception is an atomic force microscope (AFM) operated in a contact mode, where the tip of the probe is in contact with the surface of the sample with weak force and the operation is often done without feedback control.

In addition, the contact mode AFM gives a chance of detecting several interactions between tip and sample. For example, the force and the current flow can be detected simultaneously and independently with a conducting AFM probe.² If the data bits are written without topographical modification, repetition rate of data bits is not directly associated with the characteristic rate as dictated by roughness to which the AFM needs to respond. Therefore, it is possible that the rate of reading and writing the bits is not limited by the rate of AFM operation. From this point of view, one of the candidates for establishing the practical information storage system using SPM is the system constructed with a contact mode AFM and a recording medium in which the bit can be formed without topographical modification.

There have been many reports concerning the information storage with the SPM.³⁻⁸ In most cases, however, a bit is accompanied by a topographical change of the recording me-

diu, whether intentionally or not.³⁻⁵ We have already reported that a small bit of 10 nm in diameter could be formed in a polyimide (PI) Langmuir-Blodgett (LB) film by an application of a voltage pulse with the SPM probe.⁹⁻¹¹ We also showed that the bit exhibits a change in conductance only with no change in topography.^{10,11} At that time, an AFM with a conducting cantilever was used and the topographical image and current image have been observed simultaneously and independently. Hence, our measurement setup was easily adaptable for the practical information storage.

In this article, we show a feasibility study carried out to establish practical information storage with the SPM, i.e., the recognition of the bit array which represents several bits of information, and the possibility of high speed recording and reading.

II. EXPERIMENT

The PI LB film was used as the recording medium. The repeat unit of the PI consists of pyromellitic acid dianhydride-oxidianiline (PMDA-ODA). The PI LB films were deposited onto gold (111) surfaces. The detailed preparation method was described in Ref. 9. Six- or eight-layer PI LB films were used in this study. The thicknesses were 2.4 and 3.2 nm, respectively.

A home-built AFM was used and operated in the so-called variable-deflection mode, where the immediate lever deflection was recorded and no feedback control was applied to the *z* position of the medium. The spring constant of the cantilever was about 0.1 N/m and both sides were coated, about 100 nm thick, with Pt or W for acquiring electrical conductance. The cantilever was scanned with the probe tip being in contact with the surface of the recording medium. Figure 1 shows a schematic diagram of our test system. The generation of pulses, recognition of the bits from the current signal, and the scanning of the medium appropriate for writing and reading were performed under computer control.

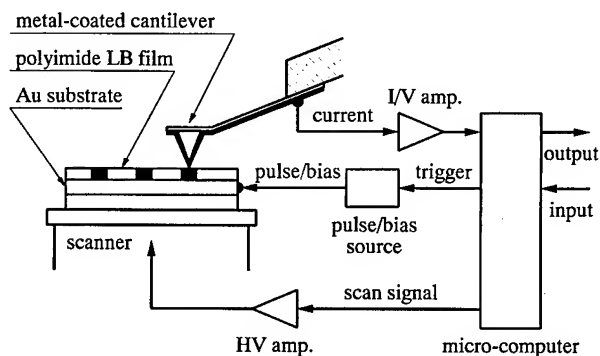


FIG. 1. Schematic diagram of a test system.

III. RESULTS AND DISCUSSION

Writing a bit pattern was carried out according to the binary coded data which represents information such as characters, text, etc. The information is in the 36×36 bit matrix. Each bit is formed by applying a $+10$ V, $5 \mu\text{s}$, rectangular voltage pulse. Figure 2(a) shows the current image of the resultant bit pattern. Figure 2(b) shows the current signal corresponding to the line indicated in Fig. 2(a). A position

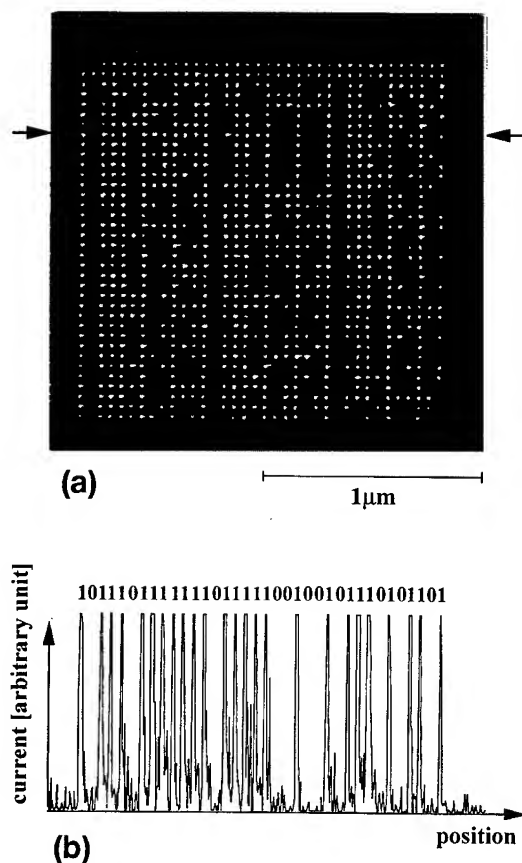


FIG. 2. Bit pattern formed after applying $+10$ V, $5 \mu\text{s}$, rectangular voltage pulses according to a set of binary data. The recording medium consisted of an eight-layer PI LB film. The area is $2 \times 2 \mu\text{m}^2$. The bias voltage applied to the medium during observation was $+2.5$ V. (a) Current image and (b) line scan profile of the current image.

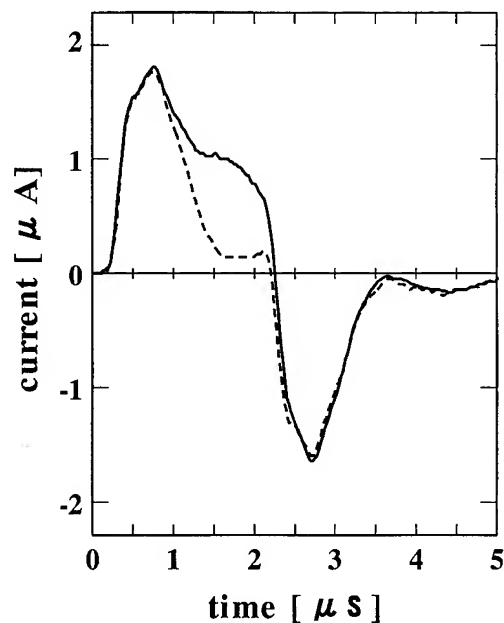


FIG. 3. Transient response of the current during application of a $2 \mu\text{s}$ wide voltage pulse. The curve shown by a continuous line is the response when the bit is formed, and the curve shown by a dashed line is that when the probe is not in contact with the medium.

where a current that was larger than the predetermined threshold was detected, was recognized as a position where a bit had been written. To find the position of the bit pattern, all bits were formed in the first two adjacent horizontal rows except the first two bits. By reading these bits out (36×2 bits), the system recognizes the position of the first bit (upper left) and the spacing of two adjacent bits. The initial bit in each horizontal array was always written to indicate the start bit. These are additional bits to those representing the information. The bit spacing serves the purpose of reducing the error rate caused by noise. The binary coded data written on the medium were well reproduced as shown in Fig. 2(b).

Figure 3 shows the transient response of current observed when the bit is formed by applying a $2 \mu\text{s}$ pulse. When the probe was not in contact with the medium, the response shown by the dashed line in Fig. 3 was observed, which seems to be caused by a stray capacitance around the probe/medium junction. Therefore, it seems that the obtained response involves the contributions of the transition of the medium, i.e., PI, to the highly conducting state and the stray capacitance. Actually, the writing with $2 \mu\text{s}$ pulses was achieved by reduction of the stray capacitance. Furthermore, the difference between two curves shown in Fig. 3 becomes obvious within $1 \mu\text{s}$. These indicate that the transition occurs whenever a voltage above a certain threshold can be applied to the junction. It may be possible to write a bit with a voltage pulse shorter than $1 \mu\text{s}$, if the stray capacitance can be reduced. A bit pattern could be also formed by applying such pulses of $2 \mu\text{s}$ width without lowering the writing probability, however a picture written under these conditions is not shown.

For faster reading, it is necessary to perform high speed

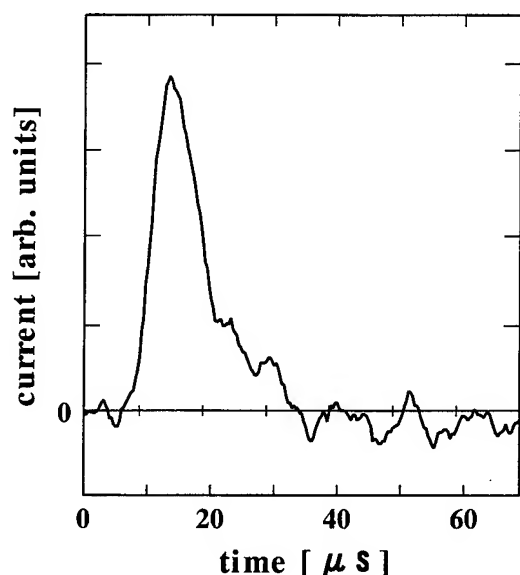


FIG. 4. Current signal when a single bit is detected using a high speed amplifier.

scanning of the probe and high speed detection of the current. In general, the current flowing between the probe and the sample is very low under SPM operation. To detect a low current at high scanning rate, we used a current amplifier designed with bipolar transistors, which is a quite general and simple circuit and has a small input capacitance. Figure 4 shows the transient response during detection of a single bit while scanning with the rate of $8 \mu\text{m/s}$. Since the current amplifier has little gain for low frequencies, only the edges of the bits are detected and Fig. 4 indicates that the edge of the bit can be detected within about $10 \mu\text{s}$. Therefore, it is expected that the reproducible detection can be performed at a rate of about 100 kbps, if the AFM probe can be scanned with higher speed. To achieve this reading rate, the next main theme is to perform the scanning at higher rate in the contact mode, which is thought to be not too difficult.

There are two ways to realize a transfer rate as high as that in present commercial storage devices. One is designing

and fabricating the cantilever with a higher resonance frequency.¹² The other is establishing the integration of cantilevers and the parallel operation of multiple cantilevers.¹³ In both cases, however, it is still essential to fabricate a sufficiently flat medium with a large area to trace a SPM probe at a high speed. LB films on an appropriate substrate seem to be a suitable recording medium in this respect.

IV. CONCLUSION

Writing a bit pattern containing information and reading the information back was performed using an AFM with a conducting probe and a LB film as recording medium in which the data bits, showing higher conductance, can be formed without topographical change. Writing of the bits was possible with a $2 \mu\text{s}$ wide voltage pulse and reading was possible within about $10 \mu\text{s}$. For practically useful performance of reading and writing with high rate, the next step is to achieve a high speed scanning of the probe.

¹C. F. Quate, NATO ASI Ser. E: Appl. Sci. **184**, 281 (1990).

²Y. Sugawara, T. Ishizaka, S. Morita, S. Imai, and N. Mikoshiba, Jpn. J. Appl. Phys. **29**, L157 (1990).

³H. J. Mamin, P. H. Guethner, and D. Ruger, Phys. Rev. Lett. **65**, 2418 (1990).

⁴H. J. Mamin and D. Ruger, Appl. Phys. Lett. **61**, 1003 (1992).

⁵S. Hosaka, T. Shintani, M. Miyamoto, A. Kikukawa, A. Hirotsune, M. Terao, M. Yoshida, K. Fujita, and S. Kammer, J. Appl. Phys. **79**, 8082 (1996).

⁶E. Betzig, J. K. Trautman, R. Wolfe, E. M. Gyorgy, P. L. Finn, M. K. Kryder, and C. -H. Chang, Appl. Phys. Lett. **61**, 142 (1992).

⁷R. C. Barrett and C. F. Quate, J. Appl. Phys. **70**, 2725 (1993).

⁸H. Kado and T. Tohda, Appl. Phys. Lett. **66**, 2961 (1995).

⁹K. Takimoto, H. Kawade, E. Kishi, K. Yano, K. Sakai, K. Hatanaka, K. Eguchi, and T. Nakagiri, Appl. Phys. Lett. **61**, 3032 (1992).

¹⁰K. Yano, M. Kyogaku, R. Kuroda, Y. Shimada, S. Shido, H. Matsuda, K. Takimoto, O. Albrecht, K. Eguchi, and T. Nakagiri, Appl. Phys. Lett. **68**, 188 (1996).

¹¹K. Yano, R. Kuroda, Y. Shimada, S. Shido, M. Kyogaku, H. Matsuda, K. Takimoto, K. Eguchi, and T. Nakagiri, J. Vac. Sci. Technol. B **14**, 1353 (1996).

¹²H. J. Mamin, L. S. Fan, S. Hoen, and D. Ruger, Sens. Actuators A **48**, 215 (1995).

¹³S. C. Minne, Ph. Flueckiger, H. T. Soh, and C. F. Quate, J. Vac. Sci. Technol. B **13**, 1380 (1995).

Semiconducting, gas-sensing properties of Europium bisphthalocyanine Langmuir–Blodgett thin films

Bingjie Liang,^{a)} Chunwei Yuan, and Yu Wei

National Laboratory of Molecular and Biomolecular Electronics, Southeast University, Nanjing, 210096, People's Republic of China

(Received 8 September 1996; accepted 26 March 1997)

In this article, Europium bisphthalocyanine (EuPc_2') ten-layer Langmuir–Blodgett films transferred to an interdigital electrode has been prepared. The LB films are fabricated microsensors for halogen gases. The sensor incorporates a microheater and a diode to control the temperature of the film. The gas concentration is measured by the conductance change. The film temperature effects on sensitivity, response and recovery time. When the sensor was operated at above 120 °C, the response had a slow downward drift, indicating the bond formation between the Cl_2 and the Pc rings which caused the ring to become chlorinated and lose its conductivity. The interaction between EuPc_2' and Cl_2 in the LB films was monitored by UV-vis spectroscopy. LB films exposed to Cl_2 gave a red-shifted Q band which is similar to the electrochemical oxidation of the macrocycle. The response time of ten-layer LB films is 45 s and the recovery time is 1 min in the vapor of 50 ppm Cl_2 in room temperature. © 1997 American Vacuum Society. [S0734-211X(97)11204-5]

I. INTRODUCTION

Phthalocyanine and many of its derivatives exhibit properties which are interesting for applications in materials science. In particular, these compounds have been proposed as new materials for gas sensors.

Lanthanide bisphthalocyanine complexes are well known for their semiconductive properties.¹ The adsorption of certain gaseous molecules on the surface of Pc crystals and thin solid films produces significant changes in the electrical properties of these materials. Gas adsorption on LB films of conjugated macrocycles seems to promote rapid changes in electrical conductivity.^{2,3}

Rare-earth-metal bisphthalocyanine are electrochromic. LB films of the material show distinctive colors changes when oxidized or reduced despite their thinness. A film which exhibits marked changes in both electrical conductivity and optical absorption in the presence of an atmosphere containing small quantities of an impurity gas would be advantageous for incorporation into devices particularly if the color change are to be sufficient for visual observation.⁴

Deposition of thin films by the LB technique is a highly attractive method for producing well-ordered thin films of known thickness. Faster response were observed with LB films substituted phthalocyanine molecules when exposed to NO_2 than those reported for vacuum evaporated film.⁵ Surface damage is inherent with high-energy techniques such as sputtering or evaporation, therefore, fabrication of molecular layers by the room temperature Langmuir–Blodgett technique is expected to produce better gas sensitivity.

Metal phthalocyanines have two possible sites for gas adsorption: one is the central metal itself and the other is the π -electron system. The chemical properties of the two sites may change considerably and may, in turn, reflect the adsorption properties of the Pc towards oxidative gases.⁶ Ce-

rium bisphthalocyanine⁷ and ytterbium bisphthalocyanine⁴ LB thin films were found to produce significant change in the electrical conductivity for the vapor of NO_x and chlorine, respectively. Wang⁸ reported that LB films of phthalocyanine $[(\text{C}_6\text{H}_{13})_3\text{SiOSiPcOGePcOH}]$ was fabricated microsensors for halogen gases, which had good selectivity for the gas to be measured.

In this article, the electrical properties of ten-layer LB film of Europium bisphthalocyanine derivative transferred to an interdigital electrode are reported, and the variation of optical absorbance in the visible region was monitored using UV-vis spectroscopy.

II. EXPERIMENT

A. Synthesis

Due to the intermolecular interaction between the macrocycles, peripherally unsubstituted metallophthalocyanines are practically insoluble in common organic solvents. The solubility of phthalocyanines can be increased by introducing substituents into the periphery of the macrocycle. These substituents lead to a larger distance between the inclined stacked phthalocyanine and enable their solution. In this article, Eu (III)-octa-4-(tetrahydrofurfuroxy) phthalocyanine was synthesized. It has good solubility with organic solvents such as chloroform. The structure of Eu bisphthalocyanine is shown in Fig. 1.

B. Description of the Interdigital Comb Electrode

Since the conductivity of Pc LB films is low, an interdigital microelectrode is required for reliable electrical measurements. The substrates used for conductance measurement of LB films were glass substrates precoved with gold interdigital electrodes. The gold interdigital electrode consists of 50 finger pair of electrodes giving a width of 50 μm , spaced

^{a)}Electronic mail: 1bj@seu.edu.cn

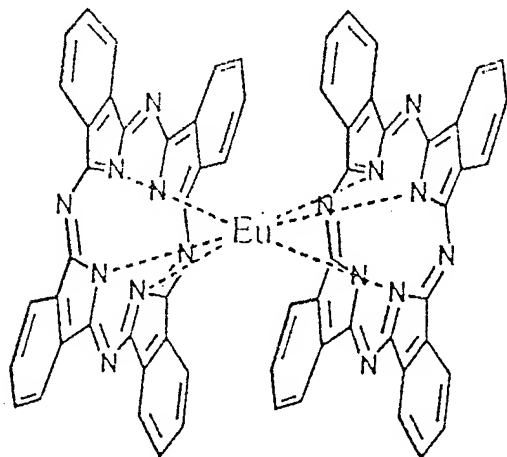


FIG. 1. The structure of Europium bisphthalocyanine.

50 μm from the adjacent electrode. The finger overlap distance was 10 mm. The structure of the gas sensor is shown in Fig. 2.

These substrates were well washed first with detergent, followed with freshly distilled water, the methanol and chloroform using ultrasonic waves for 15 min each, dried and immediately used.

C. Film preparation

Eu bisphthalocyanine was dissolved in chloroform under ultrasonic waves. A given amount of chloroform solution (5×10^{-4} M) was spreaded onto the freshly distilled water ($\sim \text{pH}$ 6, 20 $^{\circ}\text{C}$). After complete evaporation of the solvent, the surface pressure area of isotherms were measured by compressing the mobile barrier at a speed of $2 \times 10^{-3} \text{ nm}^2 \text{ mol}^{-1} \text{ s}^{-1}$.

Deposition was Z type with the substrate being passed through a clean water surface in one compartment and withdrawn through the monolayer of Eu phthalocyanine in the other compartment. The LB films were transferred to the substrate by the vertical dipping method at a surface pressure 25 mN/m, up-down speed is 5 mm/min.

For UV-vis measurements, monolayers were transferred to a Quartz disk that was cleaned in chromosulfuric acid

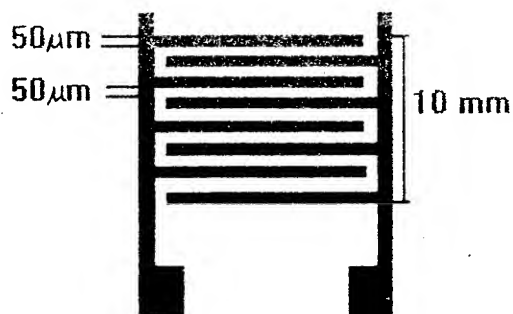


FIG. 2. Structure of gas sensor.

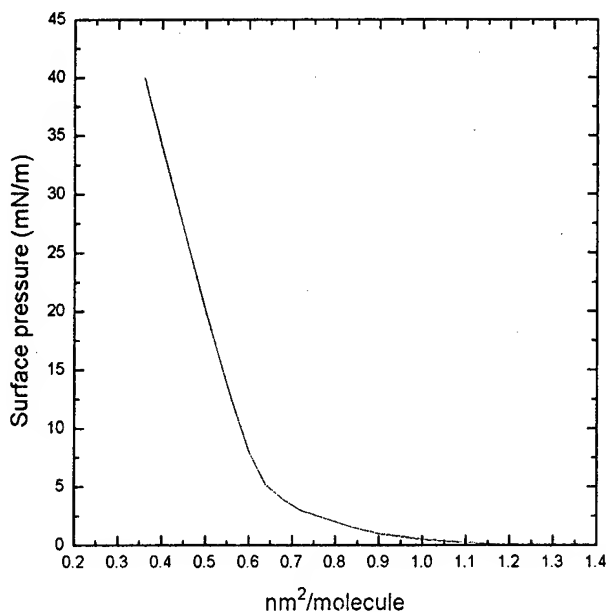


FIG. 3. The π -A isotherm of Europium bisphthalocyanine at 20 $^{\circ}\text{C}$.

maintained at 110 $^{\circ}\text{C}$. After 2 h in the cleaning solution, the slides were copiously rinsed with distilled water and placed in NaOH solution (pH 11) for 12 h. The resulting slides were hydrophilic.

D. Gas-sensing experiments

The Cl_2 used was custom blended with N_2 of ultrahigh purity. Before each chlorine test, the entire gas-sensing system was cleaned with N_2 . In a typical experiment, the sensor was heated to the desired temperature and the system was purged with the Cl_2 - N_2 gas mixture for 15 min then the Cl_2 gas was switched off and N_2 was turned on for 15 min.

E. Spectroscopic analysis

UV-vis absorption spectra were recorded on a Shimadzu 2100 UV-vis spectrophotometer.

III. RESULTS AND DISCUSSION

A. Monolayer behavior

The isotherm given in Fig. 3 shows that the monolayer can be compressed to a fully condensed close-packed film. The reproducibility of the isotherm is excellent and is limited only the accuracy with which known amounts of the Eu phthalocyanine can be applied to the trough. The collapsing pressure of the monolayer is 40 mN/m. The value for the average area per molecule obtained by extrapolating the steeply rising part of the curve to zero pressure is approximately 0.6 nm^2 . This value may be associated with a tilted edge-on molecular organization on the water surface.

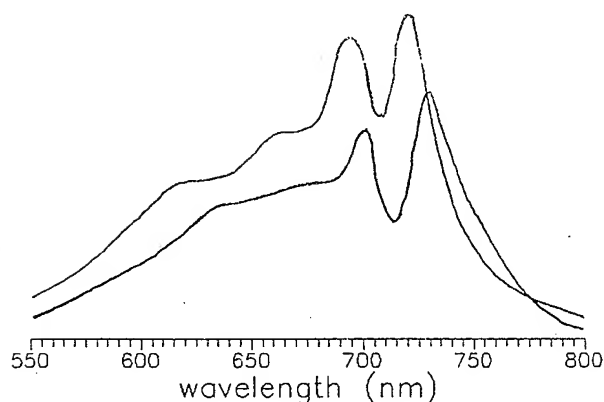


FIG. 4. The absorption spectra for ten-layer Eu bisphthalocyanine LB films.

B. Film characterization

The phthalocyanine compounds exhibit strong electronic transitions in the visible region (Q band at 600–800 nm) and in the near ultraviolet one (b or Soret band at 300–400 nm) due to π – π^* transition states of Eu symmetry. The absorption spectrum for ten-layer Eu bisphthalocyanine LB films on Quartz disc is illuminated in Fig. 4. The Q band splits into two maximum peaks at 690 and 718 nm. Also shoulder peak splits into 620 and 660 nm, respectively. The spectrum are typical in the spectrum of solid (bulk and thin solid films) lanthanide bisphthalocyanine derivatives.

Figure 5 shows how the absorbance value, measured at 690 nm, varies with the number of layers. The absorption increases linearly with the number of deposited layers. Though the molecule is not of the classical amphiphilic type,

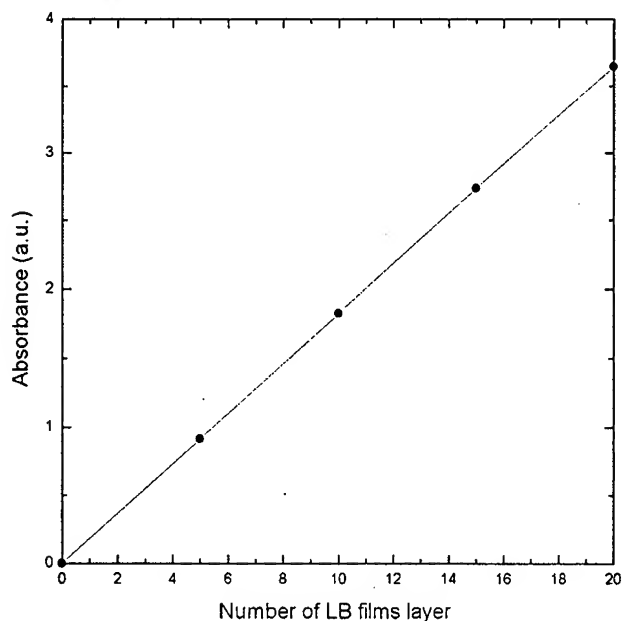


FIG. 5. Absorbance, measured at 690 nm, against number of layers.

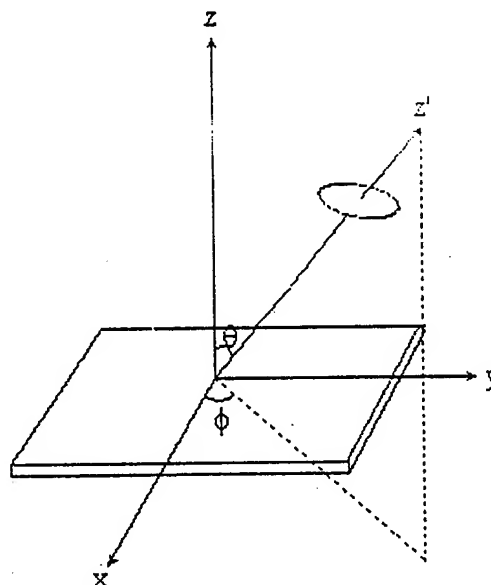


FIG. 6. Coordinates (x, y, z) for expressing Eu bisphthalocyanine orientation Y axis: dipping direction, z axis: center axis of Eu phthalocyanine represented by the circular plate.

a straight line plot is obtained indicating reproducible quality and adhere tenaciously to substrates and to each other.

The optical properties of phthalocyanine in the visible region arise from two electronic transitions of the conjugated π -electron system. The transition dipole moment of both bands lie in the plane of the phthalocyanine ring so that the polarization of the optical absorption can be used to determine the molecular orientation. The polarization plane or the incident beam was chosen perpendicular to the dipping direction and the glancing angle selected experimentally was 0° and 30° .

In order to evaluate the particular orientation of R_8YbPc_2 molecules on LB films, Cartesian coordinates (X, Y, Z) are introduced as shown in Fig. 6. The Y and Z axes are chosen to be parallel to the dipping direction and normal to the substrate, respectively. To simplify the situation, the phthalocyanine ring is considered as a flat circular plate which the π – π^* transitions dipole is uniformly distributed. The center axis of the plate, Z' , is oriented at an angle θ to the Z axis with an azimuth of ϕ in the X – Y plane. When polarized light is illuminated on the film at an angle is β' to the Z axis in the Z – X plane with a refraction angle β inside the layer (Fig. 7). Using a dichroic ratio at 690 nm, $D_0=0.82$ and $D_{30}=0.29$, the orientation angles could be calculated as $\theta=35.7^\circ$ and $\phi=57.5^\circ$.

C. Gas sensibility

Chemisorption of a gas capable of inducing donor–acceptor interaction produces an ionized state and deforms the energy band structure. The response and recovery curves of the sensor to Cl_2 (50 ppm in N_2) at various temperatures are shown in Fig. 8.

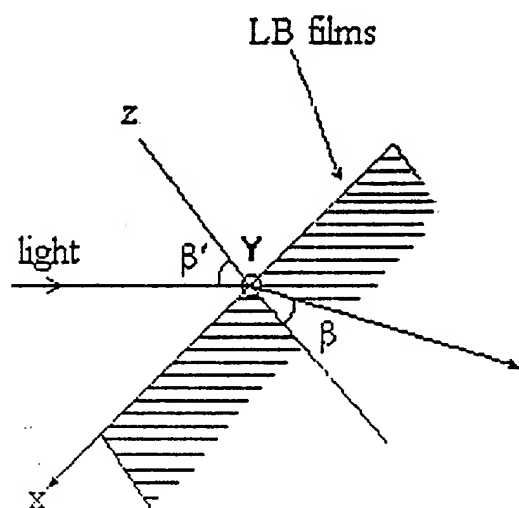


FIG. 7. Direction of polarized incident light on the substrate.

At the presence of chlorine, Cl_2 leads to an irreversible oxidation of the Pc_2Eu LB films. First, Cl_2 causes a strong increase of the conductivity, then an important conductivity decrease is observed with a concomitant transformation into $\text{Pc}_2\text{Eu}^+\text{Cl}^-$, as evidenced by the optical absorption spectrum. The peak height and the peak width decreased. The peak wavelength produced a redshift by 8 nm which is characteristic of the "oxidized form." Electrochemical studied of Lanthanide bisphthalocyanine films have also shown that redshifted Q band is produced upon oxidation.⁹

Cl_2 first induces the formation of Pc_2Eu^+ and consequently the concentration of charge carriers increases. However,

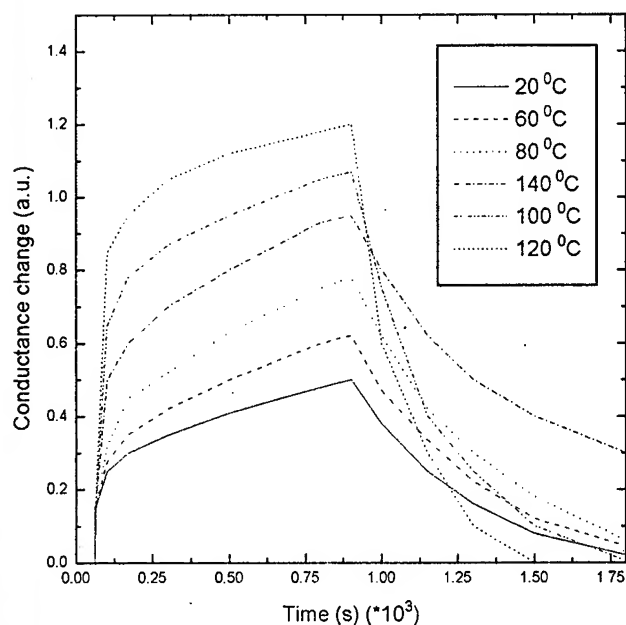


FIG. 8. The response and recovery curves to 50 ppm Cl_2 in N_2 at various temperatures.

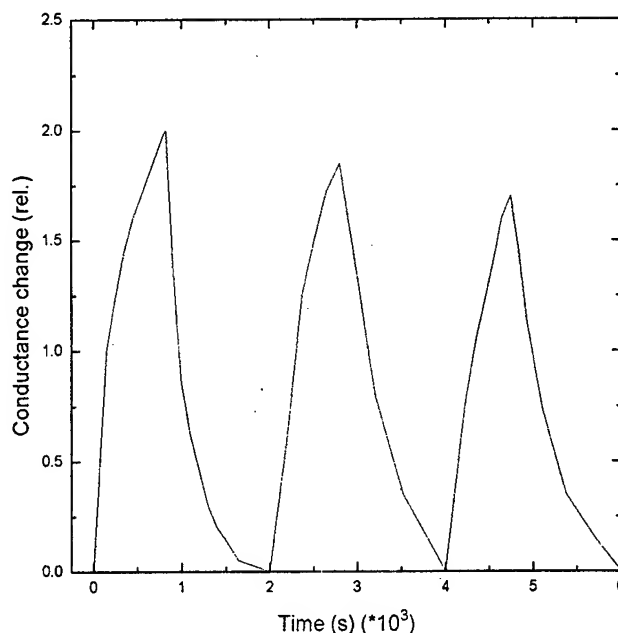
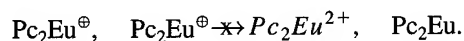


FIG. 9. Response of EuPc_2' LB films cycled at 15 min exposure to Cl_2 flowed by 15 min N_2 cleaning.

when the concentration of Pc_2Lu^+ predominates over that of Pc_2Lu , the conductivity of the thin films must decrease, since charge transport via electron exchange can no longer occur:



The conductivities of the films increased sharply with time at first, which may be due to the surface absorption effect, and the increased slowly, probably due to the bulk diffusion. The

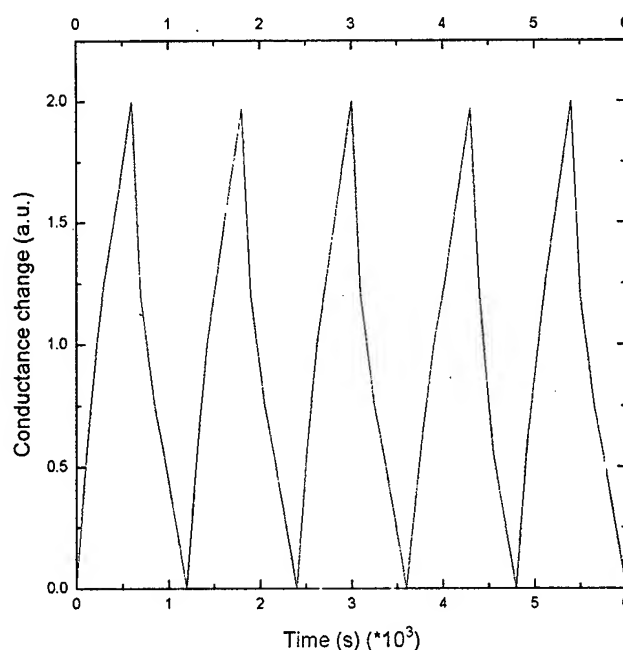


FIG. 10. Response of EuPc_2' LB films by using pulsed operation at 1 min exposure to Cl_2 .

adsorption and the desorption processes occurred simultaneously, producing a dynamic equilibrium and hence saturation in the conductance values. When the Cl_2 gas was removed, the recovery also showed a rapid decrease followed by a slower decrease of conductance. The responses increase with temperature. When operated at high temperature ($>120^\circ\text{C}$), the sensor response had a slow downward drift when Cl_2 was present, indicating the bond formation between the Cl_2 and the phthalocyanine rings which caused the ring to become chlorinated and lose its conductivity.

The reproducibility of the conductivity changes in an atmosphere containing 50 ppm Cl_2 has been studied. Figure 9 shows the degradation of the sensor response when cycled at 15 min exposure to Cl_2 of followed by 15 min N_2 cleaning. The degradation can be improved by using pulsed operation to reduce the exposure time to Cl_2 from 15 to 1 min with solenoid valve and time control. The result is shown in Fig. 10.

At room temperature, the response time is 45 s and the recovery time is 1 min. After Cl_2 were removed, the conductance does not return to the baseline. It may be attributed to its strong electronegativity and the bulk effect.

IV. CONCLUSION

Thin film of EuPc_2' formed by the Langmuir-Blodgett technique at room temperature is sensitive to Cl_2 . The phtha-

locyanine LB films have such good temperature stability and well-controlled uniformity and thickness that the device yields highly reproducible gas-sensing results. The conductance increases with temperature at the same time. The variation of optical absorbance in the visible region has been compared in the presence of N_2 and chlorine. In room temperature, the peak height and width decreased in the presence of 50 ppm chlorine.

The lanthanide bisphthalocyanine is a promising gas-sensitive material. The sensitivity can be improved by searching for a suitable central lanthanide atom and peripherally substituted groups.

¹S. Robinet and C. Clarisse, *Thin Solid Films* **170**, L51 (1989).

²W. A. Snow and W. R. Barger, in *Phthalocyanines Properties and Applications*, edited by C. C. Leznoff and A. B. P. Lever (VCH, New York, 1989).

³H. Wohltjen, W. R. Barger, A. W. Snow, and N. L. Jarvis, *IEEE Trans. Electron Devices* **ED-32**, 1170 (1985).

⁴H. Y. Wang, W. H. Ko, D. A. Batzel, M. E. Kenney, and J. B. Lando, *Sens. Actuators B* **1**, 138 (1990).

⁵A. Pauly, J. P. Blanc, S. Dogo, J. P. Germain, and C. Maleysson, *Synth. Met.* **55-57**, 3754 (1993).

⁶S. Mukhopadhyay and C. A. Hogarth, *Adv. Mater.* **6**, 162 (1994).

⁷R. Aroca, H. Bolourchi, and D. Battisti, *Langmuir* **9**, 3138 (1993).

⁸K. R. Rickwood, D. R. Lovett, B. Lukas, and J. Silver, *J. Mater. Chem.* **5**, 725 (1995).

⁹M. Petty, D. R. Lovett, and J. M. O'lonnor, *Thin Solid Films* **179**, 387 (1989).

Formation of Si nanowire by atomic manipulation with a high temperature scanning tunneling microscope

R. Hasunuma, T. Komeda,^{a)} and H. Mukaida

Joint Research Center for Atom Technology-Angstrom Technology Partnership, Higashi 1-1-4, Tsukuba, Ibaraki 305, Japan

H. Tokumoto^{b)}

Joint Research Center for Atom Technology, National Institute for Advanced Interdisciplinary Research, Higashi 1-1-4, Tsukuba, Ibaraki 305, Japan

(Received 20 February 1997; accepted 5 March 1997)

The formation of Si nanowire during indentation of the scanning tunneling microscope tip onto the Si(111) surface was investigated by changing the sample bias, the temperature, and the tip retraction speed. The wire length at room temperature is in the order of 1 nm, however, the wire was elongated remarkably with either increasing temperature or bias voltage while keeping a positive sample bias. The wire was also elongated when the tip speed was decreased. The typical length was ~ 14 nm at 481 °C, +2.0 V and 320 nm/s. In order to explain these results, we proposed a simple model for the wire formation by taking into account the surface diffusion and electromigration effects. © 1997 American Vacuum Society. [S0734-211X(97)01604-1]

I. INTRODUCTION

Characterization of structures and electric properties of nanoscale features is one of the most attractive subjects, since the novel nature would appear to be important as the ultralarge scale integrated (ULSI) devices are more scaled down. For this purpose, the scanning tunneling microscope (STM) is a powerful tool since it has the capabilities of both the observation and formation of nanoscale features in atomic scale accuracy.¹⁻⁸ In particular, the manipulation of the Si surface has attracted special interests and many articles have been published on the subject.^{1,3-7} Very recently, we have found that the Si(111)- 7×7 surface can be controllably removed in a layer-by-layer manner by approaching the biased tip until it contacts the substrate and then retracting it.¹ In this technique, the surface Si atom removal seems to occur with the field assisted and chemical reaction between the tip and surface. During this removal process, the nanoscale Si wire is formed between the tip and Si surface. When the tip is retracted further after the nanoscale contact, the nanoscale wire is elongated until breaking, at which time we define the wire length. The typical length is ~ 1 nm at room temperature as reported in Ref. 1. Here we expect that the wire length may depend on the sample bias as well as temperature, since at high biases more Si atoms were removed and at high temperature, there exists many mobile Si atoms on the surface. In this article, we present the experimental results on the formation of nanoscale wires by changing the sample bias, the temperature, and the tip retraction speed, and discuss the mechanism of the formation of nanoscale Si wires by considering the mass transport around the tip-substrate point contact region in through to the surface diffusion and electromigration.

II. EXPERIMENT

The clean 7×7 surface was prepared by flashing the $1\times 6\times 0.3$ mm³ *p*-type (B-doped, 0.01 Ω cm) Si(111) substrate up to 1100 °C. The STM tip was prepared using a poly crystalline W rod by electrochemical etching, and was cleaned by repetitive heating at 900 °C in UHV with the electron bombardment method. For high temperature experiments, the substrate was heated resistively by running the dc current across the sample. At the first stage of the manipulation, we set the tunneling sample voltage and current as 1.5 V and 0.5 nA, respectively, to fix the initial tunneling distance prior to the indentation processes. This tunneling condition was used for imaging before and after indentation. In the manipulation mode, we used a special technique already described in Ref. 1. In the present case, the feedback circuit was set off and the sample bias was tuned to appropriate values between 2.0 and -0.8 V. here the bias voltage at high temperatures was given as the same as in Ref. 1: in order to raise the sample temperature, V_{dc} was applied across the sample and the tip was approached to the nearly central point, giving a bias voltage of $V_{dc}/2$. Then the tip approached the surface, where the tip excursion distance was fixed at 0.8 nm for all present experiments. After the tip contacted the surface, it was retracted until the Si wire was broken. The tip excursion speed was varied from 40 to 960 nm/s. The wire breaking was identified by monitoring the current between the tip and the substrate during tip excursion. The current was recorded by switching a preamplifier from high gain for imaging to low gain for measuring currents as high as 100 μ A. The temperature was varied from 396 to 481 °C and was monitored with a pyrometer. The thermal drift of the system was less than 0.02 nm/s in both the *x-y* and *z* directions, which was much slower than the tip excursion speed.

^{a)}Present address: Texas Instruments Tsukuba R & D Center, Tsukuba, Ibaraki 305, Japan.

^{b)}Permanent address: Electrotechnical Laboratory, Tsukuba, Ibaraki 305, Japan.

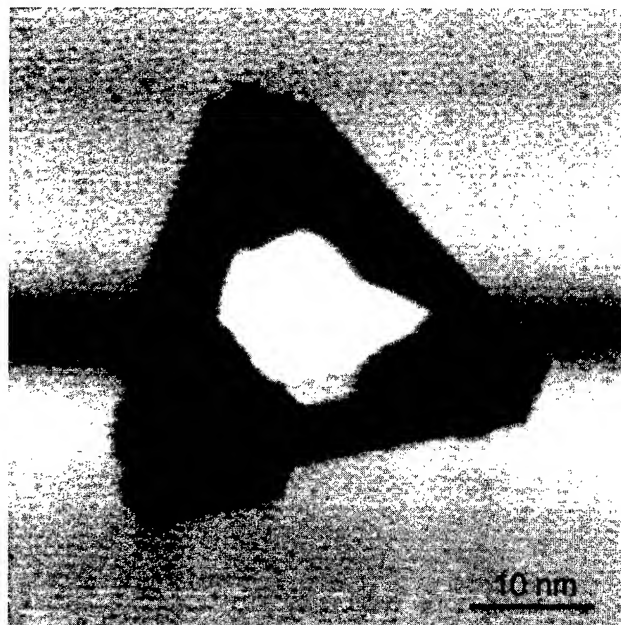


FIG. 1. A $50 \times 50 \text{ nm}^2$ STM image of Si(111)- 7×7 surface after indentation at 396°C , the sample voltage of 1.2 V, and the tip excursion speed of 320 nm/s. The hole with a depth of 0.3 nm was created by the indentation process. The clear 7×7 structure appears at the bottom of the hole. The bright protrusion in the center of the hole is the evidence of the wire formation between the tip and the surface.

III. RESULTS

Figure 1 shows a typical STM image of Si(111)- 7×7 surface after indentation at a temperature of 396°C , a sample bias of 1.2 V and a tip excursion speed of 320 nm/s. We clearly see a hole with a depth of 0.3 nm, whose size was dependent on both the sample bias and temperature. At the bottom of the hole, we see the clear 7×7 structure, at which the reconstruction occurred since it was higher than the 2×1 - 7×7 phase transition temperature.⁹ at room temperature, we see other structures such as 2×2 , c - 2×4 , and $/3 \times /3$. In the hole, there exists a protrusion which may be the result of the junction formation between the tip and surface, as frequently observed in the case of metals.¹⁰⁻¹² In other words, this can be regarded as evidence for the formation of the Si wires. The Si wire length was defined by the amount of the z -piezo motion from the contact to the point where the junction broke and was measured to be $\sim 6 \text{ nm}$ in this particular case. The wire length was dependent on the sample bias, the temperature, and the tip retraction speed. The results obtained at the tip retraction speed of 320 nm/s and various temperatures are summarized in Fig. 2. It is clearly seen that the wire length is elongated with an increase of temperature, and has a polarity dependence where the Si wire was elongated only at positive sample bias. This polarity dependence was not caused by the substrate heating induced by the current between the tip and the substrate, since it shows no rectification at an elevated temperature. The tip excursion speed dependence at 396°C and various bias voltage are summarized in Fig. 3. The wire length is elongated with a decrease in the tip excursion speed. These results indicate that the Si wire

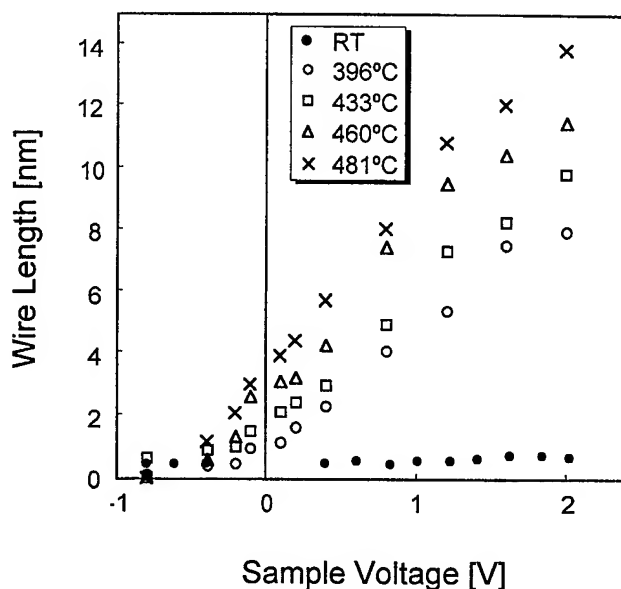


FIG. 2. The sample bias dependence of the Si wire length at the various temperatures while keeping the tip retraction speed at 320 nm/s.

length strongly correlated with how many Si atoms were supplied from the surrounding during indentation. The sample bias and temperature dependence show that the growth of the Si wire can be explained by surface diffusion of the Si atoms and electromigration due to electric field between the tip and the substrate.

IV. DISCUSSION

A. Surface diffusion

First, we discuss the Si adatoms behavior in the absence of the electric field between the tip and the substrate. The wire length at 0 V was obtained by averaging the values at 0.1 and -0.1 V . The tip speed dependence on the wire length is shown in Fig. 4 at the various temperatures. We

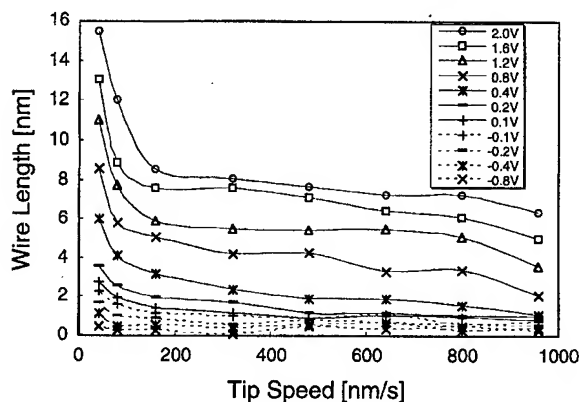


FIG. 3. The tip speed dependence of the Si wire length at the various sample bias voltages between 2.0 and -0.8 V and at 396°C .

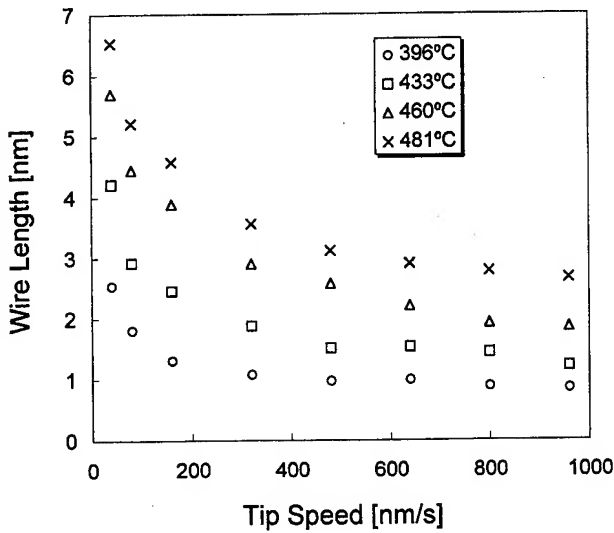


FIG. 4. The tip speed dependence of the Si wire length at 0 V obtained by averaging the values at 0.1 and -0.1 V.

have previously reported that the Si atom removal occurs by either a field assisted chemical reaction or the field evaporation at room temperature.¹ However, it can not explain the results obtained here very well, since the layer-by-layer surface removal could be extended to more than 10 nm away from the contact region at a high temperature, while it is typically less than 3 nm at room temperature. Here we assume that the Si atoms transfer could occur by the thermal diffusion process from the hole edges generated at the initial stage of the indentation. As soon as the supplement of the Si adatoms from the hole edge to the tip-substrate contact region during retraction is stopped, the Si wire is broken. To correlate the wire growth with the surface diffusion, the side wall view of the Si wire is assumed to be circular as shown in Fig. 5. This assumption may stand when the energies of the tip-wire interface and wire-substrate interface is much smaller than the surface energy of the wire. Based on this model, the volume of the Si wire with a length L is

$$V = 0.1\pi L^3. \quad (1)$$

Then,

$$dV/dt = 0.3\pi\nu L^2, \quad (2)$$

where $\nu (=dL/dt)$ is the tip retraction speed. Since the Si atoms are supplied from the edge of the hole as discussed above, the flux ϕ_D of the Si adatoms from the hole edge to the tip-substrate contact region can be given by

$$\phi_D = -2\pi LD \partial n / \partial r \, dt, \quad (3)$$

where D , n , and r are a diffusion coefficient expressed by $D = D_0 \exp(-E_{ac}/k_B T)$, surface density of mobile Si adatoms, and the distance from the tip-substrate contact center, respectively. Under the condition that the hole grows in a two-dimensional manner with a single bilayer depth in most cases, the area of the hole should be proportional to the vol-

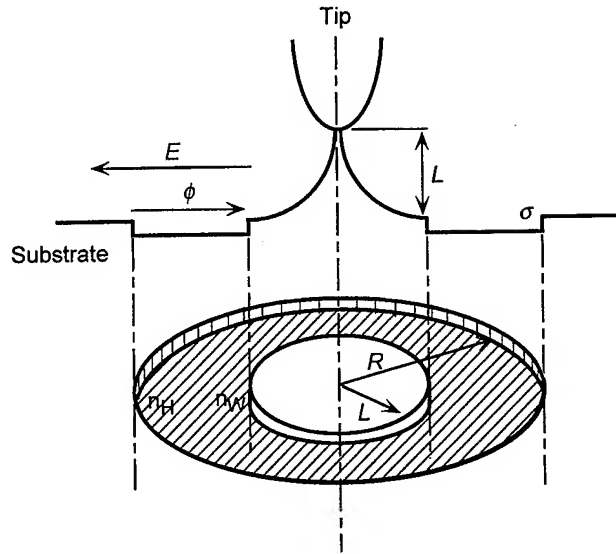


FIG. 5. A schematic drawing of a model to explain the wire growth. The side wall view of the wire is assumed to be a circular shape which is adequate when it is in the thermal equilibrium. Here L , R , n_H , n_W and σ , E and ϕ are the wire length, the distance between the hole edge and tip-substrate contact center, the density of Si adatoms at the hole edge, the density at the wire edge, the electric field induced by sample voltage, and the flux of Si atoms, respectively. In this model, the Si atoms are supplied from the hole edge, which turns out to be that the volume of the hole is equal to that of the wire.

ume of the Si wire. In addition, the quantity $\partial n / \partial r$ is inversely proportional to r because of the continuity of the flux. Then

$$\partial n / \partial r = -(1/L)(n_H - n_W) \log(0.1L/\sigma + 1)^{1/2}. \quad (4)$$

Here n_H , n_W , and σ are the density of Si adatoms at the hole edge, the Si wire edge, and the thickness of a bilayer of the Si(111) surface, respectively. Thus, at a point where the Si wire is broken during the tip retraction process from the Si surface, the flux should balance with the volume expansion as

$$\rho \, dV/dt = \phi_D, \quad (5)$$

where ρ is the volume density of Si. Then, from Eqs. (2)–(4), we obtain the following relation:

$$L^2 \log(0.1L/\sigma + 1)^{1/2} \propto D/\nu. \quad (6)$$

Figure 6 shows the Arrhenius plots of the quantity $L^2 \log(0.1L/\sigma + 1)^{1/2}$ given by Eq. (6) for several tip speeds at 0 V. From the slope of these curves, we can deduce the activation energy as $E_{act} = 1.4 \pm 0.2$ eV. This value is close to the previously reported activation energy, 1.3 eV for surface diffusion.¹³ However, the activation energy obtained from the slope of the curves in Fig. 6 slightly increases with increasing the tip speed, for which there are several possible reasons. The first is that the density of the Si adatoms at the Si wire edge is determined by the rate of volume expansion of the wire: the higher volume expansion by high speed retraction causes the less density of adatoms at the wire edge. The second is that the shape of the wire is not always circular as assumed before, which could be varied by the tip speed.

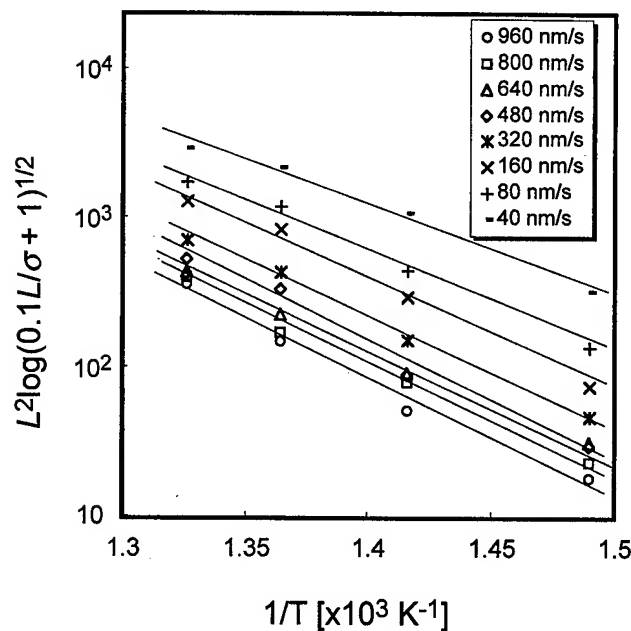


FIG. 6. Arrhenius plots of the quantity $L^2 \log(0.1L/\sigma + 1)^{1/2}$ obtained for 0 V at various tip speeds. From these slopes, we obtain the activation energy of 1.4 ± 0.2 eV.

When the tip retraction speed is faster than relaxation time of the wire, the system cannot be in thermal equilibrium. Therefore, it can be said that the activation energy of 1.2 eV obtained at the slowest speed of 40 nm/s is the most reliable under the present model.

B. Electromigration effect

When there exists an electric field induced by the sample voltage, the flux of the Si atoms from the surroundings to the wire is expressed as the sum of the thermal diffusion ϕ_D and the electromigration terms ϕ_E as

$$\phi = \phi_D + \phi_E = (-2\pi LD \partial n / \partial r - 2\pi LD n_C q Z E / k_B T) dt, \quad (7)$$

when n_C , q , Z , E , and k_B are the density of the charged Si adatoms, the elementary charge, the effective charge of the adatoms, the electric field, and a Boltzmann constant, respectively. For simplification, we assume that the local electric field on the surface is determined by surface conductance when the tip-substrate contact is formed which may stand at a point far from the tip: at a point close the tip, we should take into account the dipole moment induced by the tip.¹⁴ Since the current flows isotropically from the contact, the electric field E can be expressed as

$$E \propto 1/r. \quad (8)$$

When the tip sticks just onto the Si surface, the boundary conditions can be expressed as $V=0$ at $r=3 \times 10^{-1}$ cm, which is the distance between the edge of the substrate and tip-substrate contact center, and $V=V_S$ at $r=1 \times 10^{-7}$ cm, which is the radius of the contact. The radius has assumed from the result that the adatoms are removed at room tem-

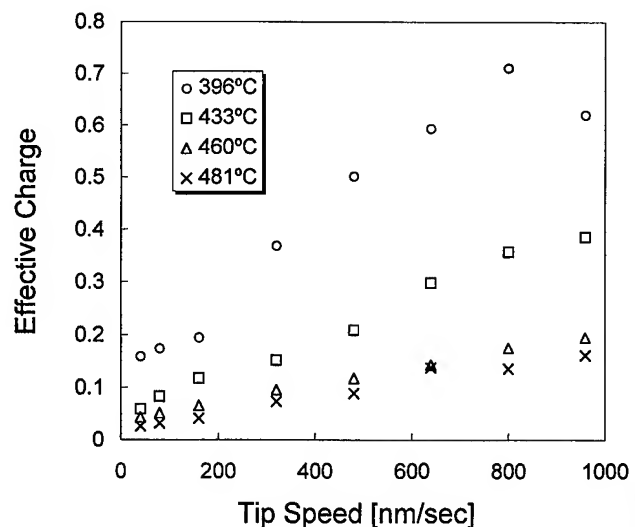


FIG. 7. The tip speed dependence of the effective charge at the various temperatures. In the case of the low tip speed and high temperature, the values are close to the previously reported value of 0.05.

perature in the area of less than 1×10^{-14} nm², which reflects the contact size. Thus, the electric field at r is

$$E = 0.06 V_S / r, \quad (9)$$

where V_S is the sample bias. From Eq. (9), the electric field at $r=1$ μm is given as 1.5×10^3 V/cm when $V_S=1$ V. Here the electric field induced by the current for substrate heating is ~ 10 V/cm which is much smaller than the field at $r < 1$ μm. It means we can ignore the heating current induced field in the present experiment. The rate of volume expansion of the Si wire is balanced with ϕ at the rupture point, then

$$0.3\pi\rho\nu L^2 = -2\pi LD \partial n / \partial r - 0.3\pi D n_S q Z V_S / k_B T, \quad (10)$$

where

$$D = \xi a^2 \omega \exp(-E_{act}/k_B T). \quad (11)$$

By assuming that the sample bias dependence of ϕ_D is negligible when the temperature is not changed, we obtain the relation from Eq. (10) as

$$L^2 = A V_S + C. \quad (12)$$

Here n_S , ξ , a , and ω are the density of mobile surface atoms on Si(111), hopping probability, the distance between nearest hopping sites, and the vibration speed of surface atoms, respectively. To simplify the discussion, here $n_C=1.5 \times 10^{15}$ cm⁻² which is the density of surface atoms on Si(111), $\xi=1$, $a=0.77$ nm which is the distance between adjacent adatoms on dimer adatom-stacking fault (DAS) structure,¹⁵ and $\nu=1 \times 10^{13}$ s⁻¹. From the slope of L^2 vs. V_S , we obtain $Z=0.03-0.71$ and summarized Z as a function of the tip speed in Fig. 7. The positive values of Z indicate that the Si atoms are positively charged at these temperature. It can be clearly seen that the effective charge is dependent on the tip speed and temperature. This is probably caused by the over-simplified model which we use: the shape of the wire is not always kept at a circular shape when the tip

speed is fast and temperature is low. Further, we can not ignore the Si atoms migration on the tip surface. For a more precise explanation, we need to observe the wire shape directly with the other experimental tools such as TEM. However, as long as the tip speed is slow and the temperature is high, the present model could stand and therefore the obtained effective charge of 0.03 may be accurate and coincide with the simulated value of 0.05.¹⁶

V. CONCLUSIONS

We demonstrated Si atom manipulation with which the surface bilayer was removed by approaching the biased tip to the Si surface. During this manipulation, the Si nanoscale wire was formed between the tip and substrate. By analyzing the wire length with a proposed model, we deduced the activation energy of 1.4 ± 0.2 eV for the thermal diffusion of Si atoms on a Si(111)- 7×7 surface. From the data for the biased tip, we found the electromigration effect in the wire formation process. Eventually we deduced the effective charge of 0.30–0.7. It is concluded that the proposed model is adequate when the tip retraction speed is slow and temperature is high enough to realize the thermal equilibrium.

ACKNOWLEDGMENT

This work was supported by NEDO.

- ¹T. Komeda, R. Hasunuma, H. Mukaida, and H. Tokumoto, *Appl. Phys. Lett.* **68**, 3482 (1996).
- ²R. S. Becker, J. A. Golovchenko, and B. S. Swartzentruber, *Nature (London)* **325**, 419 (1987).
- ³I.-W. Lyo and Ph. Avouris, *Science* **253**, 173 (1991).
- ⁴Ph. Avouris, I.-W. Lyo, and Y. Hawegawa, *J. Vac. Sci. Technol. A* **11**, 1725 (1993).
- ⁵Ph. Avouris and I.-W. Lyo, *Appl. Surf. Sci.* **60**, 426 (1992).
- ⁶H. Uchida, D. H. Huang, J. Yoshinobu, and M. Aono, *Surf. Sci.* **287**, 1056 (1993).
- ⁷H. Uchida, D. H. Huang, F. Grey, and M. Aono, *Phys. Rev. Lett.* **70**, 2040 (1993).
- ⁸D. M. Eigler and E. K. Schweizer, *Nature (London)* **344**, 524 (1990).
- ⁹R. M. Feenstra and M. A. Lutz, *Surf. Sci.* **243**, 151 (1991).
- ¹⁰L. Kuipers and J. W. Frenken, *Phys. Rev. Lett.* **70**, 3907 (1993).
- ¹¹J. I. Pascual, J. Méndez, J. Gómez-Herrero, A. M. Baró, N. García, and Vu. Thien Binh, *Phys. Rev. Lett.* **71**, 1852 (1993).
- ¹²A. Ohi, W. Mizutani, and H. Tokumoto, *J. Vac. Sci. Technol. B* **13**, 1252 (1995).
- ¹³A. V. Latyshev, A. L. Aseev, A. B. Krasilnikov, and S. I. Stenin, *Surf. Sci.* **227**, 24 (1990).
- ¹⁴L. J. Whitman, J. Stroscio, R. A. Dragoset, and R. J. Celotta, *Science* **251**, 1206 (1991).
- ¹⁵K. Takayanagi, Y. Tanishiro, M. Takahashi, and S. Takahashi, *Surf. Sci.* **164**, 367 (1985).
- ¹⁶D. Kandel and E. Kaxiras, *Phys. Rev. Lett.* **76**, 1114 (1996).

Study on photoelectric properties of a TiO₂ nanoparticle

Yongmei Chen, Yaan Cao, Yubai Bai,^{a)} Wensheng Yang, and Jihua Yang
Department of Chemistry, Jilin University, Changchun 130023, People's Republic of China

Haiyan Jin

Center of Test, Jinlin University, Changchun 130023, People's Republic of China

Tiejun Li

Department of Chemistry, Jilin University, Changchun 130023, People's Republic of China

(Received 6 March 1997; accepted 10 March 1997)

Here we report a study of surface states in nanometer-sized TiO₂ particles (*P1*) prepared in our lab with the aid of surface photovoltage spectroscopy, electric field-induced photovoltage spectroscopy (EFISPS), and fluorescence spectrum. The diameter determined by x-ray diffraction is 3.5 nm. The EFISPS of *P1* shows several subband gap absorption bands at 350, 375, and 460 nm, respectively. The fluorescence peaks of dry powder and an aqueous suspension of *P1* at 361, 391, and 469 nm were detected. The above results are concluded to be relative with the located surface states within the band gap. The surface states were deduced to be oxygen vacancies, i.e., Ti³⁺, Ti²⁺, Ti⁺, whose maximum densities lie at 0.32, 0.50, and 1.02 eV below the maximum density of states in the conduction band, respectively. © 1997 American Vacuum Society. [S0734-211X(97)01704-6]

I. INTRODUCTION

Nanometer-sized TiO₂ particles have been investigated intensively in recent years owing to their applications in solar energy conversion¹ and water detoxification.² The interfacial transfer of the photogenerated charge carriers is thought to be the essential step in these processes. Until now, theories to analyze charge transfer and recombination in TiO₂ were developed for the "ideal" interface where electrons and holes can exist only in valence or conduction bands. In fact, surface states located within the band gap may mediate charge transfer or provide recombination centers. However, it is so complicated to assign the energies, distribution, and nature of surface states that the data are usually open to various interpretations. Salvador has used transients from subband gap illumination and concluded that surface states due to Ti³⁺ and oxygen deficiencies are present near the conduction band edge.³ Luminescence experiments by other groups showed several peaks with subband gap wavelengths.⁴ These peaks were assigned to several surface states associated with species formed during the photo-oxidation of water.

Surface photovoltage spectroscopy (SPS) is a useful method to study the surface feature of a solid, while the electric field-induced photovoltage spectroscopy (EFISPS) is a newly developed method which gives more information about surface states in the semiconductor.⁵ With the aid of x-ray diffraction (XRD), SPS, EFISPS, and fluorescence spectrum, we can identify the energetic positions and features of the surface states in TiO₂ nanoparticles.

II. EXPERIMENT

A. Preparation of nanometer-sized TiO₂ powder

Nanometer-sized TiO₂ particles were prepared following the method reported by Kormann *et al.*⁶ 2.5 ml TiCl₄ was slowly added to 800 ml of 0 °C de-ionized water (18 MΩ) while vigorous stirring. The resulting transparent colloidal suspension was then dialyzed twice against de-ionized water for a total of 20 h, the final pH was 2.5. The hydrosol was concentrated to about one tenth of its original volume with the aid of a rotary evaporator (60 °C), then was dried at room temperature for several days leading to the formation of a glassy solid which was ground to obtain a white shiny powder (*P1*). This powder can be resuspended in water steadily.

The TiCl₄ used was analytical reagent grade and without further purification. The water used was distilled twice with de-ionized water with an electric resistance of 18 MΩ at room temperature.

B. Characterization

An XRD pattern was recorded by a RIGKU x-ray meter. The SPS of TiO₂ powders were obtained using an apparatus built in our lab, which is composed of a light source, a monochromator, a sample cell, a lock-in amplifier, and a computer control system. The electric field-induced SPS (EFISPSs) were performed while external fields were applied to the samples. The ITO conducting glasses were used as the transparent front and back electrodes. When the positive electrode is on the illuminated side, the applied bias is called positive. The fluorescence spectra of dry and aqueous *P1* were recorded by a Shimadzu spectrofluorophotometer RF-5000.

^{a)} Author to whom correspondence should be addressed.

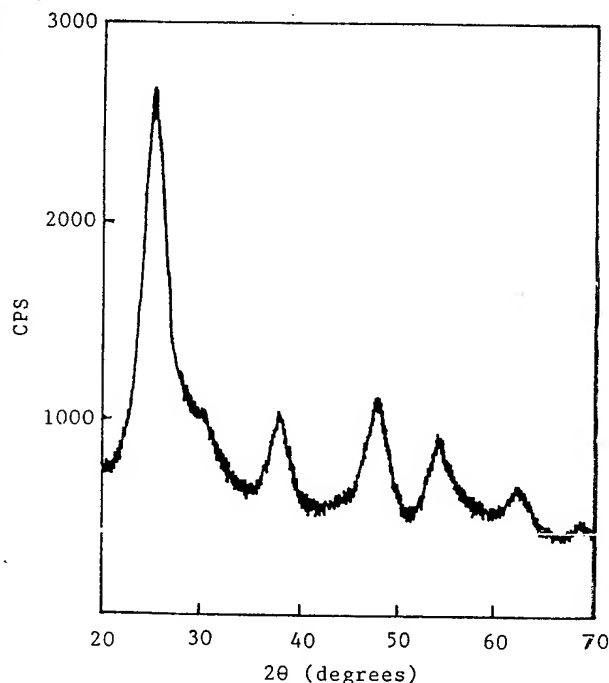


FIG. 1. X-ray diffraction pattern for P1.

III. RESULTS AND DISCUSSION

A. Character of nanometer-sized TiO₂ particles

XRD of P1 is shown in Fig. 1. The calculated lattice constants confirm that P1 is anatase TiO₂, while the diffraction profile were obviously broadened. Taken the small crystallite size as the only reason for the broadened line, crystallite size can be determined through the Scherrer equation:

$$L = K\lambda / \beta \cos \theta,$$

where θ is the Bragg angle, λ is the wavelength of the x ray, L is the mean dimension of the crystallites composing the powder, β is the breadth of the pure diffraction profile at half-maximum intensity, and K a constant approximately equal to unity and related to the crystallite shape, here taken as 0.89. The calculated crystallite size of P1 is 3.5 nm.

XRD and transmission electronic microscopy (TEM) are useful methods to determine the size of fine particles. In the case of TiO₂, the particle shape is often irregular, and sometimes aggregation occurs among particles, so precise determination of the diameter of a single particle is difficult. XRD is the method used to identify the crystallite size. The particle size determined by TEM is larger than that by XRD, which can be interpreted that the "particle" appearing in TEM is an aggregation of several crystallites.

B. SPS and EFISPS of nanometer-sized TiO₂ particles

The SPS of P1 shows a band edge at 360 nm [see Fig. 2(b)], corresponding to the band gap of 3.44 eV. The band gap is widened compared to the bulk value (3.3 eV), which is the result of the so-called quantum size effect.⁷

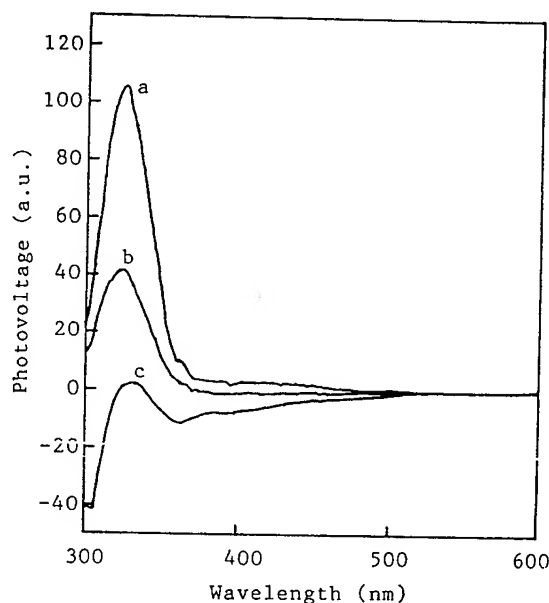


FIG. 2. Surface photovoltaic spectra of P1; (a) at negative field (-0.5 V); (b) no external field; (c) at positive field (+0.5 V).

When the external electric field was -0.5 V, the spectrum was seldom changed except for an increase in the intensity of the band with the maximum at 330 nm decreased and new absorption bands were observed at 375 and 460 nm, respectively. The EFISPS of P1 is shown in Figs. 2(a) and 2(c).

Driven by the external field, the charge carriers in surface and spatial charge region can be redistributed. If an external positive field is applied to a *p*-type material, the photovoltaic response will be increased, for the direction of the internal field in *p* type is consistent with that of the external positive field. For *n*-type material, because the internal field is opposite to the external positive field, the photovoltaic response will be decreased. Based on the principle of EFISPS and the

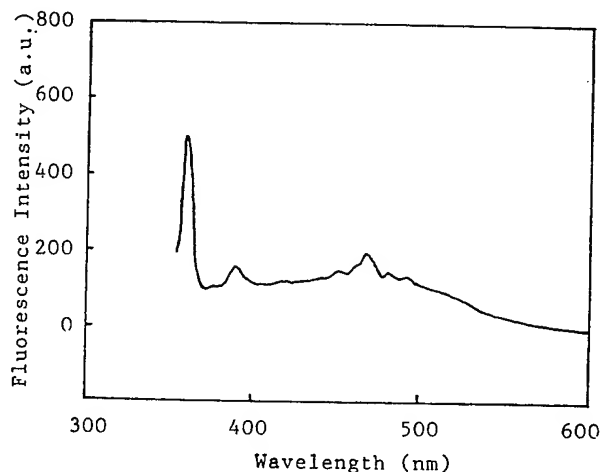


FIG. 3. Fluorescence spectrum of P1 powder measured in the air at room temperature. Excitation wavelength is 338 nm.

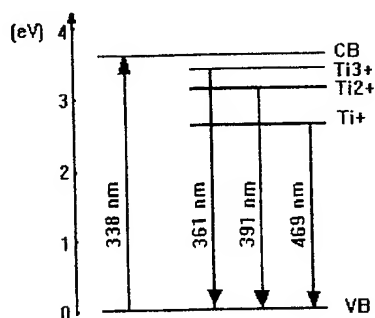


FIG. 4. The scheme of relative energy position of the conduction band, valence band, and surface states in $P1$.

results in our experiment, we identified $P1$ as n -type material.

The subband gap absorption bands were assigned to the transitions from the valence band to the surface states within the band to the surface states within the band gap. The maximum of these bands lie at 350, 375, and 460 nm, respectively.

TiO_2 always behaves as an n -type material due to oxygen vacancies in the crystal structure.⁸ At the same time the surface states in $P1$ are assumed to be oxygen vacancies, so Ti^{3+} , Ti^{2+} , and Ti^+ were produced, respectively. That is, one oxygen vacancy leads to 2Ti^{3+} , missing rows of oxygen from several Ti^{2+} , and metallic Ti exposed to the surface is said to be Ti^+ . Considering the feature of these surface states, we assigned bands 350, 375, and 460 nm to Ti^{3+} , Ti^{2+} , and Ti^+ , respectively. In order to precisely determine the energetic position of these surface states, we studied the fluorescence of $P1$.

C. Fluorescence spectra of nanometer-sized TiO_2 particles

The fluorescence spectrum of $P1$ powder is shown in Fig. 3. With excitation of 338 nm, there are several emission peaks which appeared with the maximum at 361, 391, and 469 nm, respectively. The fluorescence of $P1$ suspended in water showed almost the same emission spectrum with that of $P1$ powder. The energetic positions of these surface states are consistent with the value determined by EFISPS, which means that the ground state level corresponded to valence band, the excited state corresponded to the conduction band, and the others are surface states, shown in Fig. 4.

Other groups have observed fluorescence in TiO_2 .^{4,9} The strong fluorescence observed in our experiment can be interpreted as follows: $P1$ was prepared under controlled condi-

tions. Each initial particle was a new-formed nucleus or poor-grown particle of the nucleus, which had large structural defects and high surface energy, so aggregation among these initial particles (3.5 nm) and the interaction among them was some kind of short-distance interaction, i.e., chemical bonding, oxygen bridge, etc. However, this kind of linkage could not be as regular as the crystal structure. The structural defects, i.e., oxygen vacancies, existed inside the particle's trap electron with energy level located within the band gap, i.e., the surface states. Moreover, because of the short distance interaction, the quencher, which is water in our experiment, could not go into the inside of the particle to trap the photogenerated electron, so radiation transition from the surface states to the valence band occurred.

IV. CONCLUSION

The nanometer-sized TiO_2 particle ($P1$) prepared in our experiment proved to be an aggregation of several initial particles through some kind of chemical bonding. The combined particles manifesting quantum size effect is the result of the confinement within the initial particle (3.5 nm by XRD). The structure defects, i.e., oxygen vacancies of the combined particles surface states, can be detected by EFISPS. The structural defects of the initial particles still exist in the combined one. Because of chemical bonding among the initial particles, the fluorescence quencher could not be reached at these defects, so strong fluorescence can be observed. Because the EFISPS and fluorescence spectrum resulted from the same resource, the energetic positions of surface states deduced from both methods is consistent with each other. From the above results, the feature and the energetic position of surface states of $P1$ were determined.

¹B. O'Regan and M. Gratzel, *Nature (London)* **353**, 737 (1991).

²W. Choi, A. Termin, and M. R. Hoffmann, *J. Phys. Chem.* **98**, 13669 (1994).

³P. Salvador, *Surf. Sci.* **192**, 36 (1987).

⁴Y. Nakato, H. Ogawa, K. Morita, and H. Tsubomura, *J. Phys. Chem.* **90**, 6210 (1986); G. Nogami, Y. Ogawa, and Y. Nidhiyama, *J. Electrochem. Soc.* **135**, 3008 (1988); G. Nogami, R. Shiratsuchi, and S. Ohkubo, *ibid.* **138**, 751 (1991).

⁵D. J. Wang, W. Liu, H. Tian, Z. Q. Zhu, G. Y. Hong, and T. J. Li, *International Symposium on Hydrogen System, Beijing* (Pergamon, Oxford), 1985, Vol. 1, p. 255.

⁶C. Kormann, D. W. Bahnemann, and M. R. Hoffmann, *J. Phys. Chem.* **92**, 5196 (1988).

⁷U. Koch, A. Fojtik, H. Weller, and A. Henglein, *Chem. Phys. Lett.* **122**, 507 (1985).

⁸W. Gopel, J. A. Anderson, D. Frankel, M. Jeahing, K. Phillips, J. A. Schafer, and G. Rucker, *Surf. Sci.* **139**, 333 (1984).

⁹(a) L. Forss and M. Schubnell, *Appl. Phys. B* **56**, 363 (1993); S. K. Poznyak, V. V. Sviridov, A. I. Kulak, and M. P. Samtsov, *J. Electroanal. Chem.* **340**, 73 (1992); M. Anpo, M. Tomonari, and M. A. Fox, *J. Phys. Chem.* **93**, 7300 (1989).

Formation of nanocrystals in *a*-Si thin films induced by pulsed laser ultraviolet irradiation

Z. Chvoj,^{a)} V. Cháb, and O. Borusík

Institute of Physics AVCR, Cukrovarnická 10, 162 00 Praha 6, Czech Republic

(Received 8 September 1996; accepted 4 April 1997)

We extend the theoretical model of recrystallization of *a*-Si films induced by pulsed laser irradiation and determine c-Si cluster distribution in the solid phase and we discuss the possibility of nanocluster formation in the *a*-Si thin layer. Our theory is based on the description of nucleation and growth of the crystalline phase in molten Si during fast heating and cooling processes. We are able to describe nonequilibrium solidification, including the distribution of nanocrystals as a function of laser pulse energy. Calculations of temperature distribution during recrystallization processes and the kinetics of crystalline phase formation show that nucleation of c-Si in *l*-Si must be taken into account in simulations of recrystallization of *a*-Si induced by high energy density pulsed laser irradiation ($>200 \text{ mJ/cm}^2$). The presence of nanoclusters in the Si layer is possible if recrystallization starts from the strong supercooled *l*-Si, e.g., if *a*-Si is melted immediately.

© 1997 American Vacuum Society. [S0734-211X(97)05004-X]

I. INTRODUCTION

Extremely nonequilibrium processing during ultraviolet (UV) laser annealing (XeCl and ArF excimer lasers) of thin films can result in a significant modification of surface properties. Pulsed laser-induced recrystallization of *a*-Si is one possible technology by which to prepare nanocrystalline c-Si.¹ Crystalline nanoclusters in thin Si films may be responsible for the emission of visible blue light during laser illumination; thus the formation of such clusters in thin Si films could result in materials that would be promising for important technical applications.

To describe the crystallization processes in *a*-Si and the conditions that lead to nanocluster formation, the dynamics of phase transformations have to be studied. Usually the recrystallization processes of *a*-Si are described by the numerical solution of a moving boundary problem involving heat conduction with a volume heat source term (see for example, Refs. 2–4). Application of this procedure can describe the rate of solidification, time of melting, and temperature field in the sample. The next step in the theoretical treatment of laser-induced *a*-Si recrystallization is extension of this model to the volume solidification (homogeneous nucleation) of the molten layer.^{5–9} This was first done in Ref. 3, where the kinetics of the solidification process in *a*-Si thin films induced by excimer laser pulses was studied both experimentally and theoretically. Independent of the wavelength of applied laser radiation a strong minimum was observed in the total time of the recrystallization process t_r versus energy density of pulses. Theoretical models in Ref. 3 focus on two different paths of solidification, one with and one without nucleation. But in the model with nucleation only the fraction of the crystalline phase in the volume is considered. Nucleation in *a*-Si and *l*-Si was thoroughly described in Ref. 6. In Ref. 6 models of the homogeneous nucleation and growth rate of c-Si in *a*-Si or *l*-Si dependent on temperature are summarized. The basic parameters of these processes

were determined to obtain good agreement with experimental data. The suitability of stationary and nonstationary models was discussed in respect to the usual temperatures of the sample. In this article we focus on determining the nuclei distribution in real Si layers with respect to the following facts that were determined by previous studies.

- (1) A description of crystallization by solution of the heat conductivity equation with a moving boundary between phases (*l*-Si \leftrightarrow c-Si or *a*-Si \leftrightarrow c-Si) cannot explain all the experimental observations.² In order to explain the non-monotonic dependence (with a strong minimum) of the time-of-recrystallization process on the energy density of the laser pulse, nucleation and recoalescence processes must be taken into account.³
- (2) Crystallization of *a*-Si with nanosecond pulses is a process where phase transformation occurs under nonequilibrium conditions. The melting point of *a*-Si is about 200 K lower than the melting point of c-Si and thus, during melting of *a*-Si, supercooled liquid Si appears in the system. In this case nucleation of c-Si is very probable.⁴
- (3) The formation of nanoclusters (recently experimentally detected and reported in Refs. 10–12) is connected with nucleation. The distribution of clusters as a function of temperature, time, and volume should be determined.

Now we present the calculations of cluster distribution as a function of volume, time, and temperature history (not determined in Ref. 3) for conditions that correspond to actual processes of recrystallization of *a*-Si thin films irradiated by laser pulses with various densities of energy. We discuss the possibility of nanocluster formation under these real conditions.

II. MODEL OF THE RECRYSTALLIZATION PROCESS

The distribution of clusters of c-Si in the system will be determined in the following model. Since stable clusters

^{a)}Electronic mail: chvoj@fzu.cz

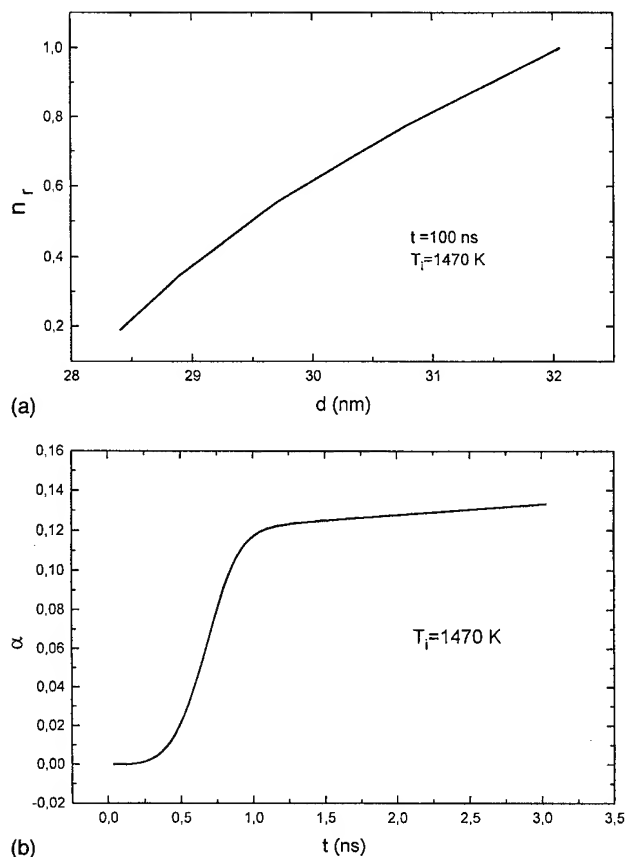


FIG. 1. (a) Relative number of clusters n_r as a function of diameter d at time $t = 100$ ns. The initial temperature is near the melting temperature of a-Si. (b) Relative volume of solid phase α as a function of time t . The initial temperature is near the melting temperature of a-Si.

form in the system as a result of a homogeneous nucleation process, it means that the first stable clusters that appear in the system are critical nuclei that relate to the instantaneous temperature. The diameters of such nuclei increase with the velocity v , which depends on the temperature. Thus at every time t we have to calculate the instantaneous diameter $d(t)$ of all clusters created during time interval (t_0, t) (t_0 is the initial time at which a-Si starts to melt). Let d_{t_i} be the diameter of clusters which appear in the system at time interval $(t_i, t_i + \Delta t)$. The number of such clusters per unit volume (assuming the temperature during small time interval Δt to be constant) is

$$N_{t_i} = j_{ss}[T(t_i)] \Delta t, \quad (1)$$

and their diameter at time t is

$$d_{t_i}(t) = 2 \int_{t_i}^t v[T(\tau)] d\tau + d^*[T(t_i)], \quad (2)$$

where $j_{ss}(T)$ is the stationary rate of nucleation, d^* is the diameter of critical nuclei at temperature T , and v is the linear growth rate of c-Si. During these calculations we must compute the real temperature in liquid Si as determined by the heat conductivity, the adsorption of laser beam energy, and the latent heat of phase transformation. This problem

was solved in Refs. 2, 3 and 13. On the basis of relations (1) and (2) we can calculate the distribution of clusters $n(d, t)$ as a function of diameter and time:

$$n(d, t) = \sum_{\tau}^* N_{\tau}. \quad (3)$$

The sum Σ^* is over the values N_{τ} which correspond to the number of all the nuclei of diameter d at time t with respect to the real courses of temperature. For simplicity we present the relative value (on a given interval of d) of distribution of clusters n_r :

$$n_r(d, t) = n(d, t) / \max\{n(d, t)\}. \quad (4)$$

The nucleation rate $j_{ss}(T)$ and growth rate v dependence on temperature were determined in Ref. 6; $\max\{n(d, t)\}$ is the maximum number of clusters at time t regardless of the diameter d . Since these values are complicated functions of material parameters (latent heat, temperature of melting, interface energy, energy barriers and kinetic parameters for diffusion and liquid-solid phase transition, change of entropy, and chemical potential), here we present only the results of their temperature dependence in Figs. 3 and 4.

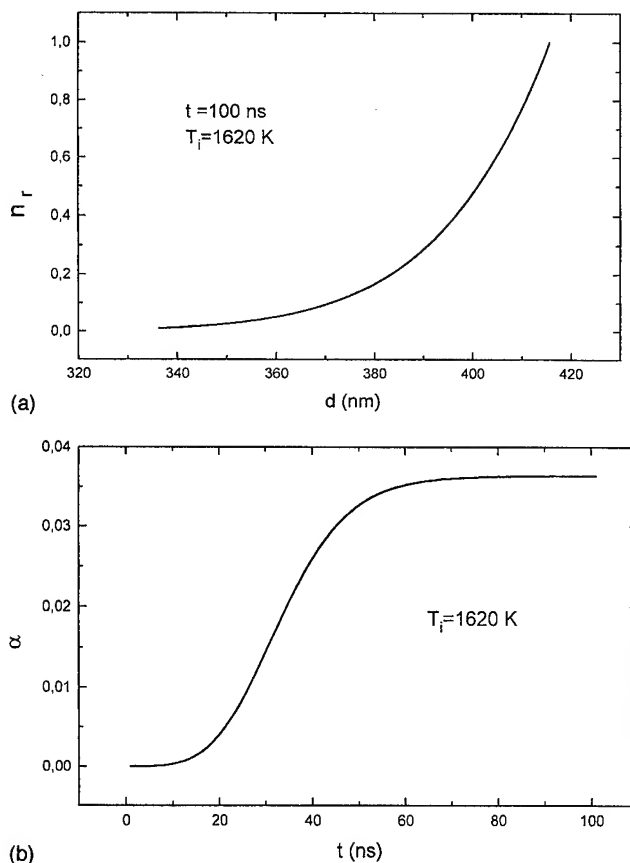


FIG. 2. (a) Relative number of clusters n_r as a function of diameter d at time $t = 100$ ns. The initial temperature is near the melting temperature of c-Si. (b) Relative volume of solid phase α as a function of time t . The initial temperature is near the melting temperature of c-Si.

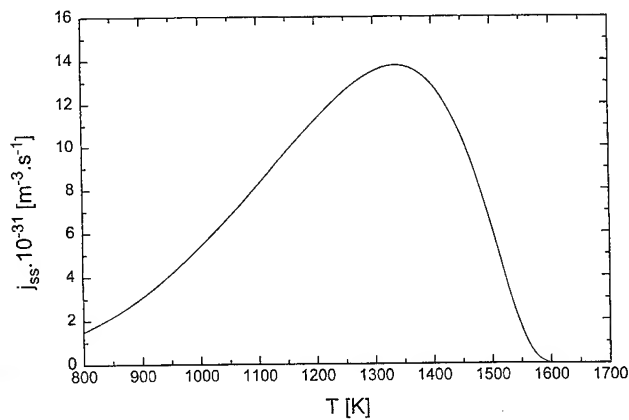


FIG. 3. The temperature dependence of the stationary nucleation rate j_{ss} for the l -Si \rightarrow c -Si phase transition.

The time evolution of the volume fraction of crystalline phase $\alpha(t) \in (0,1)$ is described by the Johnson–Mehl–Avrami–Kolmogorov equation⁴

$$\alpha(t) = 1 - \exp \left[- \gamma \int_{t_0}^t \left(\int_{\tau}^t v(t') dt' \right) j_{ss}(\tau) d\tau \right], \quad (5)$$

where γ is the geometrical factor. The value α describes the overall evolution of the crystalline phase in liquid Si during the nucleation process. Owing to the serious numerical problems, the solidification of Si resulting from the motion of the l -Si– c -Si interface is not taken into account. This problem will be the subject of a subsequent study. Here, we assume that the nuclei form and grow in the free small volume up to the maximum time of molten Si existence in order to estimate the distribution of frozen c -Si clusters in the solid phase. Subsequently the rest of l -Si solidifies at p -Si.

III. NUMERICAL RESULTS

Calculations of temperature distribution and its time evolution during recrystallization of a -Si were published earlier.^{2,3} These investigations showed that the temperature in the liquid phase of Si is practically constant in volume and is close to the melting temperature of a -Si. As soon as nucleation starts, the temperature increases very quickly to the equilibrium temperature of crystallization. Conduction of heat away from the Si layer into the substrate corresponds to an average cooling rate $T = 10^9$ deg/s. We calculated the distribution of clusters [Eq. (3)] as a function of time t , diameter, and the real course of temperature for various regimes of a -Si crystallization. Time $t = 100$ ns corresponds to the experimentally determined time of molten Si and it is supposed that, after this time, the article sample is solid phase. On the basis of these calculations we can distinguish two different regimes of crystallization.

- (1) Nucleation from high supercooled l -Si (melted a -Si, initial temperature of l -Si $T_i = 1470$ K [Figs. 1(a) and 1(b)]. The existence of nanoclusters (diameter < 100 nm) is possible in this sample, but their total volume is low in comparison with the volume of the rest of the solid

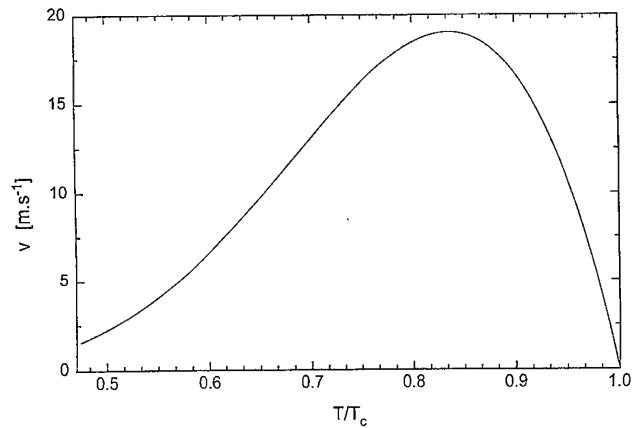


FIG. 4. The temperature dependence of the growth rate on the flat phase interface v for the l -Si \rightarrow c -Si phase transition.

phase [reflected by the low value of α ; compare Figs. 1(a) and 1(b)]. Under our conditions the rest of the solid phase is p -Si with large grains. The strong increase of solid phase volume α after 1 ns [Fig. 1(b)] is connected to the increase of nucleation. The subsequent release of latent heat increases the temperature of the l -Si and the rate of solidification substantially decreases. This process can be realized for a high density of laser pulse energy > 200 mJ/cm² when the molten layer is sufficiently large and the system has time for nucleation. In this case the transition a -Si \rightarrow l -Si \rightarrow c -Si proceeds.

- (2) Nucleation from melted pc -Si (polycrystalline Si) (initial temperature $T_i \geq 1620$ K) [Figs. 2(a) and 2(b)]. At this regime nanoclusters practically do not exist in the system; the material has a polycrystalline character. The course of function $\alpha(t)$ is similar to that in case (1), but the fast increase is realized after 30 ns. It is the result of the worse conditions for nucleation. This case corresponds to the transition (a -Si \rightarrow) pc -Si \rightarrow l -Si \rightarrow c -Si.

In Fig. 5 the “optimal” distribution of nanoclusters in the Si layer is shown, but the temperature regime leading to this distribution is hypothetical, the initial temperature of 1000 K

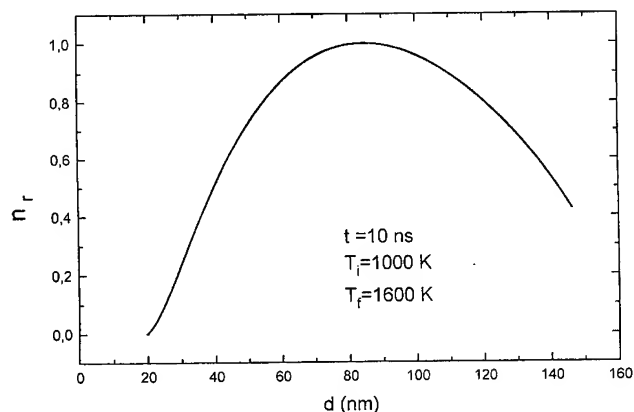


FIG. 5. The optimal hypothetical relative distribution of clusters under special regime of temperature changes.

changes to the final temperature of 1600 K in 10 ns. Thus the process of recrystallization should begin from higher supercooling (similar theoretical estimations were presented in Refs. 11 and 12). These conditions are not realistic during pulsed excimer laser irradiation of *a*-Si at room temperature.

IV. CONCLUSIONS

On the basis of our calculations we can conclude the following.

- (1) Nucleation of c-Si in *l*-Si plays an important role and must be taken into account in simulations of recrystallization of *a*-Si induced by high energy density pulsed laser irradiation ($>200 \text{ mJ/cm}^2$). This nucleation seems to be responsible for the nonmonotonic dependence of time-of-melting on the energy density of the laser pulse.
- (2) The presence of nanoclusters in the Si layer is possible if recrystallization starts from the strong supercooled *l*-Si, e.g., if *a*-Si is melted immediately. This situation occurs for higher densities of laser pulse energy.

In contrast to earlier models of *a*-Si recrystallization kinetics, our model has a broader application with more precisely defined conditions and predicts the resulting phases very well. Theoretical calculations agree well with experimental results of solidification of thin *a*-Si films during pulsed laser UV irradiation.²⁻⁴ Under certain conditions visible photoluminescence was observed,¹ which indicates the presence of nanoclusters in solid Si. The intensity of photoluminescence increases with the number of pulses. It is very important to

understand the photoluminescence mechanism, but that was not the focus of this article. We suggest that pulsed laser-induced regrowth can provide a new technique for fabricating *a*-Si thin film transistors and a "dry" method for the preparation of photoluminescent Si films.

ACKNOWLEDGMENT

The work was supported by the Grant Agency of the Academy of Sciences of the Czech Republic under Grant No. 110433.

¹R. Černý, V. Vydra, P. Přikryl, L. Ulrych, J. Kočka, K. M. A. El-Kader, Z. Chvoj, and V. Cháb, *Appl. Surf. Sci.* **86**, 359 (1995).

²R. Černý, P. Přikryl, K. M. A. El-Kader, and V. Cháb, *Int. J. Thermophys.* **16**, 841 (1995).

³O. Borusík, R. Černý, P. Přikryl, K. M. A. El-Kader, I. Ulrych, Z. Chvoj, and V. Cháb, *Appl. Surf. Sci.* **109/110**, 317 (1997).

⁴Z. Chvoj, V. Cháb, and O. Borusík, *Thermochim. Acta* **280/281**, 261 (1996).

⁵*Kinetic Phase Diagrams*, edited by Z. Chvoj, J. Šesták, and A. Trška (Elsevier, Amsterdam, 1991).

⁶Z. Chvoj, *J. Non-Equilib. Thermodyn.* **18**, 201 (1993).

⁷*Handbook of Crystal Growth*, edited by D. T. J. Hurle (North-Holland, Amsterdam, 1993), Vol. 1, p. 601.

⁸H. Kumomi and T. Yonehara, *J. Appl. Phys.* **75**, 2884 (1994).

⁹J. B. Zeldovich, *Acta Physicochim. USSR* **18**, 1 (1943).

¹⁰J. S. Im, H. J. Kim, and M. O. Thompson, *Appl. Phys. Lett.* **63**, 1969 (1993).

¹¹S. R. Stiffler, M. O. Thompson, and P. S. Peercy, *Phys. Rev. Lett.* **60**, 25 (1988).

¹²S. R. Stiffler, M. O. Thompson, and P. S. Peercy, *Phys. Rev. B* **43**, 9851 (1991).

¹³I. Lukeš, R. Šášik, and R. Černý, *Appl. Phys. A* **54**, 327 (1992).

Atomic force microscopy investigations on the surface topographies of aluminum-based composite containing nanocluster diamond

Q. Ouyang, B. Wang, and K. Okada^{a)}

Faculty of Engineering, Yamanashi University, Kofu 400, Japan

(Received 8 September 1996; accepted 19 May 1997)

Atomic force microscopy (AFM) has been used to study the rolling effect of ultrafine particles of cluster diamond contained in an aluminum-based composite during a sliding friction process. By analyzing the interaction between the tip of AFM and the spherical particles of diamond, it is estimated that at constant force mode of AFM when a particle is made to rotate, it shows a tendency to leave from the tip and a certain kind of deformed topographies are yielded. Some special images observed are explained by the mechanism based on the rotation of nanoball form particles of cluster diamond during the tip of AFM scanning over the sample surface. © 1997 American Vacuum Society. [S0734-211X(97)13404-7]

I. INTRODUCTION

Even for the most excellent solid lubricants such as graphite or MoS₂, the friction coefficient μ cannot be made smaller than 0.01.¹ On the other hand, the friction coefficient of rolling friction is about one hundred times lower than that of sliding.² If the rolling friction mechanism can be introduced into a sliding process, a sharp reduction of friction may be made. Recently, it has been expected that C₆₀ with the perfect spherical shape and weak intermolecular bonding may roll like a tiny ball bearing in a sliding contact.³⁻⁶ Friction tests have proven the lubrication effect of C₆₀, but the mechanism is still not clear.

Cluster diamond is thought to have a similar structure as C₆₀.⁷ Friction properties of aluminum-based composites containing cluster diamond have also been investigated by friction tests under a variety of conditions.^{8,9} A lowest friction coefficient, μ of 0.005, was obtained with the composite of 50 wt % diamond content at an applied load of 0.005 N. On the basis of the results, it was supposed that a certain degree of rolling friction due to the rotation of free diamond particles during friction test exists between the two contacting surfaces.

Atomic force microscopy (AFM) has been proved to be a very powerful tool in tribological studies.^{10,11} In AFM, a tip which is only a few atoms wide at the end of it, is brought into contact with a sample and is scanned across the surface. The two-dimensional real time images can be acquired by measuring the deflection signal as the tip is scanned (constant height mode), or by adjusting the height of the sample using feedback to maintain a constant deflection signal (constant force mode). In any case, the operations of AFM are based on the interacting force between the tip and the sample surface. Therefore, from the view of tribology, the AFM process can be considered as a sliding friction process between the tip and the sample. If there exist some movable parts on the sample surface, the normal working process of the whole system of AFM will be disrupted, and distorted images which do not reflect the real surface topography may be ob-

served. We thought AFM was an ideal method to study this rolling effect to see if there exists tiny spherical particles.

We report here the details of the investigation on the rolling effect of ultrafine particles of cluster diamond by AFM.

II. EXPERIMENT

The sample used was an aluminum-based composite with 50 wt % cluster diamond content prepared by a powder metallurgy. Before AFM observation, the sample was polished with emery paper (Al₂O₃, grade 3000, 5 μ m average diam), followed by the friction test and finally cleaned by acetone¹⁰ to get the friction surface.

A digital instruments Nanoscope II was utilized in the AFM observation. The cantilever was made of Si₃N₄ with a spring constant of 0.12 N/M. Scan frequency was set at 4.73 Hz which corresponds to an interval of 2.1 nm in one row of scan to get the clear image. Constant force mode was used. The normal load applied to the sample was controlled in such a way that a repelling force was always acting in the tip. Thus a positive pressure under which a particle may be moved easily was kept in the sample surface.

III. RESULTS AND DISCUSSION

Figure 1(a) shows a three-dimensional AFM image in an area of 1600 \times 1600 nm². A lot of convexes can be seen on the surface. The convexes are almost in a near rectangular shape with two sides perpendicular to the scan direction which could be produced during the friction test. An interesting observation is that on an average the slopes of these two sides are different. A typical cross section at the position indicated by the arrow in Fig. 1(a) is expressed in Fig. 1(b). The right-hand side, P, with the tip uphill, is steeper than that of the left side, Q, a downhill with the tip. The height of the sides is about 30 nm, being the total length of four or five particle diameters. There is no reason to think that a common friction test can make surface structures like this. We consider that it is not the real surface topography but a distortion of it.

^{a)}Electronic mail: okada@ccn.yamanashi.ac.jp

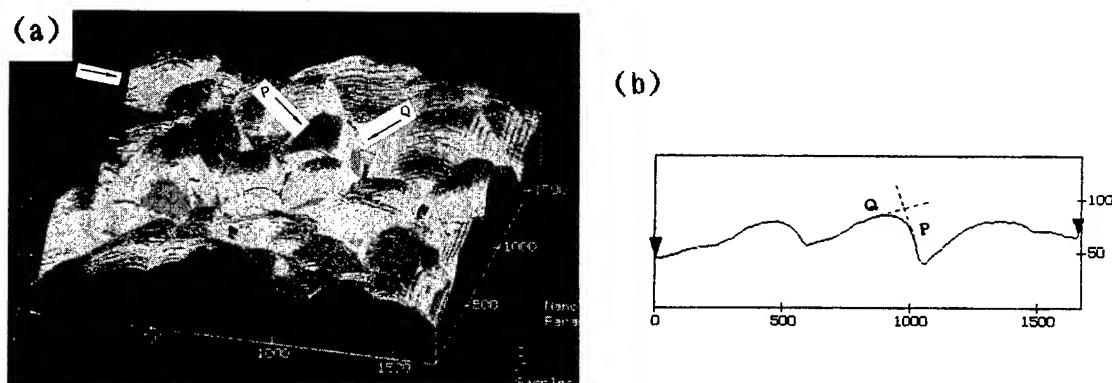


FIG. 1. (a) A $1600 \times 1600 \text{ nm}^2$ image showing different inclination on the two sides of convexes indicated by P and Q, respectively. (b) A cross section at position represented by the arrow in (a).

In addition, in an area of $400 \times 400 \text{ nm}^2$, a cliff-form structure was also observed as shown in Fig. 2(a). The wall of the cliff is parallel to the scan direction and perpendicular to the sample surface. Figure 2(b) presents the cross section of a cliff along the arrow in Fig. 2(a), the cliff is represented by A (bottom of the cliff) and B (top of the cliff). The height jumps suddenly from point A to point B by about 40 nm within about 4.2 nm along the x -scanning direction. The angle between the straight line of A, B, and the x direction is about 86° . Since the sample is not a single crystal, stepping form deformation like this is unusual. Therefore it cannot be thought to be the real surface topography.

To explain the structures mentioned above, first we take notice of the composition and structure of the sample. Unlike aluminum recrystallized into larger crystals, diamond particles are not influenced by the manufacturing process. Therefore diamond particles existing in the sample have no chemical combination with surrounding materials, which means that the interaction forces of diamond particles with surrounding materials are weaker than those of aluminum. At this condition after the friction test, the ultrafine particles which relate with each other much looser are exposed on the surface even though the free particles have been washed out by acetone. Second let us consider the forces acting on a single particle. The force that it may make the particle to

move under AFM performance is that of the tip against the sample surface, which is about a few nano Newton both in normal and lateral directions.¹² On the other hand, in order to make the particle move, it should overcome the forces resisting the gravitation of the particle and the interacting force between the particle and the surrounding materials. The weight of a particle is only at the magnitude of 10^{-12} nN , in comparison with the force acting on the particle by the tip, it can almost be neglected. As there is no chemical bond between the particle and other materials, the force acting on the particle by surrounding materials is at least no larger than that acting on this same particle by the tip, both forces are van der Waals interaction in the main.¹³ It is believed that at some contacting positions the ultrafine particles of diamond may be moved by the tip. Due to the spherical shape of the particles, rolling is much easier than sliding. On the basis of this consideration, we suppose that ultrafine particles of diamond can be made to rotate during the tip scanning over the sample surface resulting in the special images we have observed.

To simplify the problem, we analyze the movement of a single particle. Obviously the rolling direction of the particle must be along the acting force. When the scan direction is $+x$, the possible principal rolling forms of the particle are

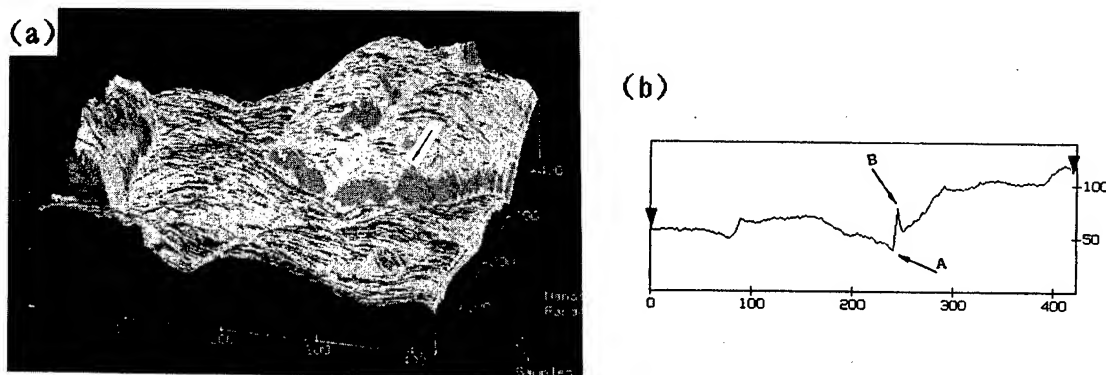


FIG. 2. (a) A $400 \times 400 \text{ nm}^2$ image with cliff-form structures parallel to the scan direction. (b) A cross section of a cliff at the position indicated by the arrow in (a). The horizontal and vertical distances between A and B are 4.22 nm and 38.53 nm, respectively.

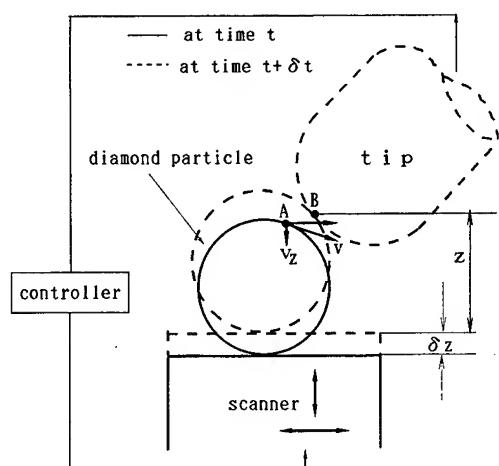


FIG. 3. A model describing the generation of the distorted topography when the tip is scanned across a diamond particle. Solid line and dot line express the relative positions of the tip and the particle at time t and $t + \delta t$, separately. At time t , contacting point A has a tangential speed v due to the rotation of the particle. To balance V_z , at time $t + \delta t$, contacting point B, the scanner makes an additive displacement δz .

clockwise in the x - z plane and counterclockwise in the x - y plane.

Figure 3 depicts a model of the creation of the distorted image caused by the rolling of a diamond particle. If the particle is rolling at time t with a tangential velocity v , at contacting point A, because vz is in the same direction as the z component of the scan of the tip, the result of the rotation is that at time t the particle shows a tendency of leaving away from the tip. Under the constant force mode we used, to stabilize the force, at time $t + \delta t$, when the tip is scanned to point B, the feedback controller will command the scanner to make a negative displacement δz , an additive value to the real height z . In this way a distorted image is created.

When the relative movement of the tip is climbing a slope, if the particles rotate in x - z plane, the negative displacement will make the slope more gentle than the real topography. On the contrary, if the tip is across down a slope, the observed topography will become steeper. The model of the effect is expressed in Fig. 4(a). When a particle rotates in x - y plane, the rolling torque is changed with the contacting position of the tip on the particle. For example, if the tip is scanned just right over the particle, the torque will become zero. Therefore the rolling of the particle will be stopped suddenly at some position when the torque is not big enough to move the particle. Figure 4(b) is a corresponding model of section trace based on the rolling of a particle in the x - y plane. Before point B, upon the same mechanism described above, the inclination of the slope shown is larger than the real one. Until point B, at which the displayed height becomes lower than the real one by $-\Delta z$, the rolling stopped suddenly. At the next row of scan, the displayed height begins to restore to the real one, h , at point C, while the difference of height with respect to point B increases to about Δz . In this way, within about 4.2 nm interval in the y direc-

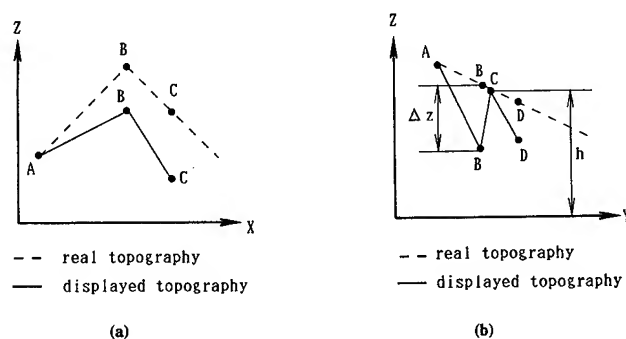


FIG. 4. The models of distorted topographies explained by the rolling effect of diamond particles when (a) particles rotate in the z - x plane and (b) when the particles rotate in the x - y plane. Dot lines express the real topography, solid lines are images observed by AFM: A, B, etc. represent the same positions on the surface.

tion, a few 10 nm of height change occurs, resulting in a cliff-form structure of Fig. 2(a).

Because the size of one particle is too small, the distortion of surface topography due to the rotation of a single particle is difficult to be identified although it exists. However the special surface structures presented here taken by AFM are thought to originate from the rotation of a few adjacent particles successively.

IV. CONCLUSION

In conclusion, the special phenomena observed by AFM can be explained by the rotation of diamond particles under the atomic force acting on it. The present results indicate that the very low friction coefficient obtained in general friction experiments may also be caused by this nanoball bearings of ultrafine particles of diamond.

ACKNOWLEDGMENTS

The authors would like to thank both the Takahashi Foundation of Industry and Economics for financial support and the Monbusho's Grand-in-Aid for JSPS Fellows.

¹K. Okada and O. Sekiguchi, J. Jpn. Soc. Precision Engineering (in Japanese) **61**, 171 (1995).

²K. Matsubara, *Tribology* (Sangyo Tosho, Tokyo, 1985) (in Japanese).

³J. H. Holloway, E. G. Hope, R. Taylor, G. J. Langley, A. G. Avent, T. J. Dennis, J. P. Hare, H. W. Kroto, and D. R. M. Walton, J. Chem. Commun. 966 (1991).

⁴B. Bhushan, B. K. Gupta, G. W. V. Cleef, C. Capp, and J. V. Coe, Appl. Phys. Lett. **62**, 3253 (1993).

⁵P. J. Blau and C. E. Haberman, Thin Solid Films **219**, 129 (1992).

⁶B. K. Gupta and B. Bhushan, Lubr. Eng. **50**, 524 (1994).

⁷Q. Ouyang and K. Okada, Appl. Surf. Sci. **78**, 309 (1994).

⁸Q. Ouyang and K. Okada, J. Vac. Sci. Technol. A **12**, 2577 (1994).

⁹Q. Ouyang and K. Okada, Trans. Jpn. Soc. Mech. Eng. (in Japanese) **61C**, 297 (1995).

¹⁰R. Luthi, E. Meyer, H. Haefke, L. Howald, W. Gutmannsbauer, and H.-J. Guntherodt, Science **266**, 1979 (1994).

¹¹A. J. de Boef, Rev. Sci. Instrum. **62**, 88 (1991).

¹²S. Grafstro, J. Ackermann, T. Hagen, R. Neumann, and O. Probst, J. Vac. Sci. Technol. B **12**, 1559 (1994).

¹³Y. Arai, *Material Chemistry of Powder Materials* (Baifu-Kan, Tokyo, 1987) (in Japanese).

Calculation of electronic energy levels in artificially confined cavities of a sphere and a circular cylinder

Yong Fang Zhao,^{a)} Xiao Gong Jing,^{b)} Lin Song Li,^{c)} Li Jun Wang,^{c)} Zheng Hui,^{c)} and Tie Jin Li,^{c),d)}

Jilin University, Changchun 130023, People's Republic of China

(Received 8 September 1996; accepted 1 April 1997)

Based on the results of experiments of the nanostructured materials, the electronic structures of the spherical and circular cylindrical well in nanometer scale are calculated using a constant potential by direct solution of the Schrödinger equation. According to the result of the bound states existing only under certain conditions, the calculated results show that the energy level structures of these modes are quite distinct from that of the hydrogen atom. The number of energy levels is finite and it depends strongly on the potential parameter. The order of energy levels is $E_{1s} < E_{2p} < E_{3d} < E_{2s} < E_{3p} < E_{4d} < \dots$ in the case of the artificially confined cavity of the sphere. It was exactly consistent with an investigation reported by T. Inoshita *et al.* [Phys. Rev. Lett. **57**, 2560 (1986)]. The calculated results of the circular cylindrical wells have similar properties on the level structure compared with that of a square spherical artificial atom well. © 1997 American Vacuum Society. [S0734-211X(97)02604-8]

I. INTRODUCTION

The technological and scientific interest in materials fabricated on a nanometer size is widespread, impacting research and development in many disciplines because of their potential utilization in electronic, optical, and micromechanical devices.¹⁻³ In addition, these materials provide an opportunity to study quantum effects not observed in bulk materials. The charge and energy of a sufficiently small particle of metal or semiconductor are quantized just like those of an atom. The current through such a quantum dot or one electron transistor reveals atomlike features in a spectacular way.⁴ For example, a novel method has been found for confining surface state electrons within artificial structures on the nanometer length scale reported by Crommie *et al.*⁵ They used a scanning tunneling microscope to move 48 individual iron atoms on a Cu (111) surface to form a "quantum corral" of radius 71.3 Å. Tunneling spectroscopy revealed a series of discrete resonances in the local density of states, an indication of size quantization; the results are close to those expected for an electron trapped in a round two-dimensional box with hard walls. On the other hand, Masuda *et al.*⁶ reported a highly ordered metal nanohole array (platinum and gold) was fabricated by a two-step replication of the honeycomb structure of anodic porous alumina. The metal hole array of the film has a uniform, closely packed array of columnar hexagonal cells with central, cylindrical, uniformly sized holes ~70 nm in diameter and from 1 to 3 μm thick. The preparation of thin metallic membranes containing uniform, patterned voids with diameters as small as 40 nm also reported by Pearson *et al.*⁷

All of these reports show that spherelike, circular cylin-

derlike or corral shaped enclosed structures were widely built in nanometer scale. So it is important to study the electron properties of these structures. It is also necessary to understand the electronic energy states of such artificial systems. Compared with experimental results, there are less theoretical studies on such structures. Recently, studies showed that in the systems of quantum confinement, the motion of electrons is governed by quantum mechanics rather than classical mechanics. Wu *et al.*⁸ studied the electron wave motion in a quantum wire with periodic structure by directly solving the Schrödinger equation; Crommie *et al.*⁵ carried out the calculation of the Schrödinger equation with "hard-wall enclosure," obtained the eigen-energies of electrons in two dimensional quantum corrals on a metal surface, and made comparisons with the experimental result. In the case of the study on the electronic structure of quantum dots, researchers have not focused on the electronic energy level in an artificially confined circular cylinder. Most research has concentrated on the electronic energy levels of the spherical artificial-like atom. For example, Hsiao *et al.*⁹ reported that the Whittaker function and the scattering Coulomb wave function are used to study the quantum confinement effect on the impurity states of a spherical confining system, especially Inoshita *et al.*,¹⁰ who performed a self-consistent local density functional calculation on the electronic structure of the superatom by using the potential consisting of band offset, ionized donors, Hartree, and exchange-correlation part.

The purpose of this article is to clarify the electronic level structure of the sphere and circular cylinder artificial atom through the calculation of electron motion in a square spherical well and a circular cylindrical well. We use the method of immediately solving the Schrödinger equation with a constant potential to calculate these two modes. The results indicate that different diameters in the nanometer scale square spherical well possess different distinct energy level structures. The different distinct energy level structures are also

^{a)}Institute of Atomic and Molecular Physics.

^{b)}Department of Physics.

^{c)}Department of Chemistry.

^{d)}Electronic mail: tjli@mail.jlu.edu.cn

obtained in calculating the circular cylinder well. Hence, these calculation will provide a basis for designing practical single electron devices. Our research is distinct from the above, one in which the calculation of the electronic structure of a single electron Schrödinger equation in a confined cavity, electron motion can be controlled by an applied potential, so there are some intriguing application backgrounds;⁵⁻⁷ second, the potential we used can be modulated, it accord with the artificial atom properties. Moreover, since the potential is simplified, the calculation becomes immediate and convenient for designing experiments, and the electronics of the spherical artificial-like atom are consistent with the investigation by a complicated potential calculation method.¹⁰

II. MODEL AND EIGEN EQUATION

The artificial-like atom is assumed to be a square spherical well or a circular cylindrical well. An electron is confined in the well. The electron interacts with a potential of the following form:

For a spherical well of the radius R_0 ,

$$V(r) = \begin{cases} -V_0 & r < R_0 \\ 0 & r > R_0 \end{cases} \quad (1)$$

For a circular cylindrical well of height l_0 and radius R_0 ,

$$V(r) = \begin{cases} -V_0 & \rho < R_0, 0 < h < l_0 \\ 0 & \rho > R_0, h < 0, h > l_0 \end{cases} \quad (2)$$

Because the potential is negative, the electron suffers an interaction of the face directing to the center. The question can be regarded as electronic motion in a square spherical or a circular cylindrical potential well. The bottom of the potential well is V_0 , the width is R_0 and l_0 , and $V_0 R_0^2$ and $V_0 l_0^2$ indicate the strength of the potential well.

The solutions to the radial part of the Schrödinger equation with this potential in spherical polar coordinates are spherical Bessel functions. The logarithmic-derivative equation that satisfied the boundary conditions at $r=R_0$ is (in a.u.)

$$\frac{1}{h_l^{(+)}(i\alpha r)} \cdot \frac{d}{dr} [h_l^{(+)}(i\alpha r)]_{r=R_0} = \frac{1}{j_l(kr)} \cdot \frac{d}{dr} [j_l(kr)]_{r=R_0}, \quad (3)$$

where $k = (E + V_0)^{1/2}$, $-V_0 < E < 0$ for bound state, and $\alpha = |E|^{1/2}$. The eigenequations obtained from Eq. (3) are

$$l=0: \quad X \cot X = -Y, \quad X^2 + Y^2 = V_0 R_0^2, \quad (4)$$

$$l=1: \quad (1/X) \cot X - 1/X^2 = (1+Y)/Y^2, \\ X^2 + Y^2 = V_0 R_0^2, \quad (5)$$

$$l=2: \quad [9 - 4X^2 - (9X - X^3) \cot X] / (3 - X^2 - 3X \cot X) \\ = (9 + 9Y + 4Y^2 + Y^3) / (3 + 3Y + Y^2), \\ X^2 + Y^2 = V_0 R_0^2, \quad (6)$$

where $X = kR_0$ and $Y = \alpha R_0$. Equations (4)–(6) are super-equations and can only be solved numerically.

For a circular cylindrical-like atom, the separation of variables of Schrödinger equation in circular cylindrical coordinates leads to two equations: one equation containing the Z variable may be solved with normal quantum mechanics methods. The eigen-equation that satisfied the boundary conditions, whose wave function and first-order derivative are continuous at $Z=0$ and $Z=l_0$, is

$$2XY \cot X - (X^2 - Y^2) = 0, \quad X^2 + Y^2 = V_0 l_0^2, \quad (7)$$

where

$$X = Kl_0, \quad K = (V_0 + E_z)^{1/2},$$

$$-V_0 < E_z < 0; \quad Y = kl_0, \quad k = |E_z|^{1/2}$$

for the bound state.

Other equation in the XY plane, the solutions of its radical part, are Bessel functions. The logarithmic derivative at $\rho = R_0$ is

$$\frac{1}{J_m(K\rho)} \frac{d}{d\rho} [J_m(K\rho)]_{\rho=R_0} = \frac{1}{K_m(k\rho)} \frac{d}{d\rho} [K_m(k\rho)]_{\rho=R_0}, \quad (8)$$

where $K = (E_{xy} + V_0)^{1/2}$, $-V_0 < E_{xy} < 0$ for bound state, and $k = |E_{xy}|^{1/2}$. Corresponding eigen-equations are

$$m=0, \quad XK_0(Y)J_1(X) = YJ_0(X)K_1(Y), \\ X^2 + Y^2 = R_0^2 V_0, \quad (9)$$

$$m>0, \quad XK_m(Y)J_{m-1}(X) = -YJ_m(X)K_{m-1}(Y), \\ X^2 + Y^2 = R_0^2 V_0, \quad (10)$$

where $X = KR_0$ and $Y = kR_0$. Equations (7)–(10) only can be solved numerically for the same reason. Finally, total energy $E = E_{xy} + E_z$.

III. RESULTS AND DISCUSSION

A. Square spherical artificial atom well

R_0 takes both values of 5 and 10 nm in this calculation. Energy levels calculated with different V_0 values are given in Table I.

(1) For $l=0$, it is obvious from Eqs. (4) that the minimum of Y is at $\cot X=0$, namely $X=\pi/2$, therefore, the condition that there is at least the bound state in this artificial atom is $V_0 R_0^2 > \pi^2/4$. Otherwise, it cannot contain an electron. The critical value of V_0 for $l=0$ is about 3.76 meV for $R_0=5$ nm and 0.94 meV for $R_0=10$ nm. We can see from Table I that, if there is no bound state for $l=0$, then there is also none for $l>0$. Therefore, if V_0 is controlled by an applied voltage, then it will be possible to realize the control of the charge in artificial atoms. This characteristic provides a probability that may become the candidate of future computers.¹¹

(2) If we use the symbols in atomic physics to indicate states of one electron, namely,

TABLE I. The calculated energy levels of the artificial square spherical atom (in meV).

V_0	n	$R_0=5\text{ nm}$			$R_0=10\text{ nm}$		
		$-E_{ns}$	$-E_{np}$	$-E_{nd}$	$-E_{ns}$	$-E_{np}$	$-E_{nd}$
0.01	1	2.9723			7.4094		
	2				0.6152	4.7896	
	3						1.6432
0.05	1	39.226			46.8248		
	2	9.4066	29.2226		37.374	43.5144	
	3			14.7592	21.9622	30.9942	39.3989
	4				2.1638	12.878	23.7924
	5						3.443
0.10	1	88.1373			96.6664		
	2	53.3695	75.8352		86.6934	93.1842	
	3	2.3069	30.5899	60.4485	70.1789	79.9123	88.7938
	4			6.2851	47.3578	60.1867	72.2003
	5				18.9279	34.3766	49.3069
	6					4.1159	20.7085
0.50	1	486.498			496.4392		
	2	446.0729	472.3896		485.7597	492.716	
	3	379.0003	418.5583	454.5964	467.97	478.4756	488.0171
	4	285.8887	338.3201	387.2388	443.0853	457.1341	470.1697
	5	168.1735	232.5594	293.9203	411.1294	428.7085	445.2643
	6	31.5355	103.7098	175.8687	372.1372	393.2271	413.2959
	7			37.8822	326.161	350.7323	374.2931
	8				273.2808	301.288	328.3054
	9				213.6288	244.9953	275.4108
	10				147.4517	182.0292	215.738
	11				75.3452	112.7513	149.5262
	12				1.2964	38.2835	77.3435
	13						2.4486

l 0 1 2 3 ...,
symbol s p d f ...,

and different energy levels with the same l value are indicated by a integer n starting from $n=l+1$, then the order of energy levels is quite different from that of the hydrogen atom. In normal atoms, the order of energy levels is

$$E_{1s} < E_{2s} < E_{2p} < E_{3s} < E_{3p} < E_{3d} < \dots,$$

while for this artificial atom it is

$$E_{1s} < E_{2p} < E_{3d} < E_{2s} < E_{3p} < E_{4d} < E_{3s} < E_{4p} < E_{5d} < \dots$$

It shall be emphasized that this sequence of energy levels completely agrees with that obtained using the self-consistent local-density function.¹⁰ With the help of Ref. 10, we would like to explain the different sequences of energy levels. From Fig. 1 we can clearly see that the s , p , d , states occupy distinct energy levels. The normal atoms are characterized by the singularly deep Coulomb potential at $r=0$, with the consequence that low angular momentum states, having a larger amplitude around $r=0$, are favored. Hence $2s$ comes between $1s$ and $2p$, $3s$ and $3p$ between $2p$ and $3d$, and so on. A square spherical well, on the other hand, lacks such a singularity, and therefore states with fewer radial nodes are favored. This property probably enables d electrons to participate in quantum transport processes in preference to $2s$, $3s$, and $3p$ electrons.

(3) The number, N , of bound states is finite in the artificial atom which is different from the case of the hydrogen atom. The number strongly depends on R_0 and V_0 values. For $l=0$ and $R_0=5\text{ nm}$, for example, $N(0.0038)=1$, $N(0.5)=6$, and $N(1.0)=8$ (the values in parentheses are V_0 , in eV).

(4) If we refer to E_{ns} , $E_{(n+1)p}$, and $E_{(n+2)d}$ as a set, the separations of energy levels in the set of lower-lying energy levels is small, while it is larger in the set of higher-lying

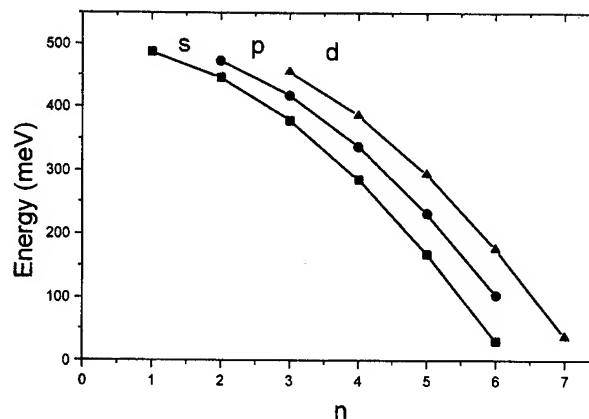


FIG. 1. Calculated orbital energies of the square spherical artificial atom well ($V_0=0.5\text{ eV}$), $R_0=5\text{ nm}$ as a function of n . The three curves represent s , p , d states, respectively.

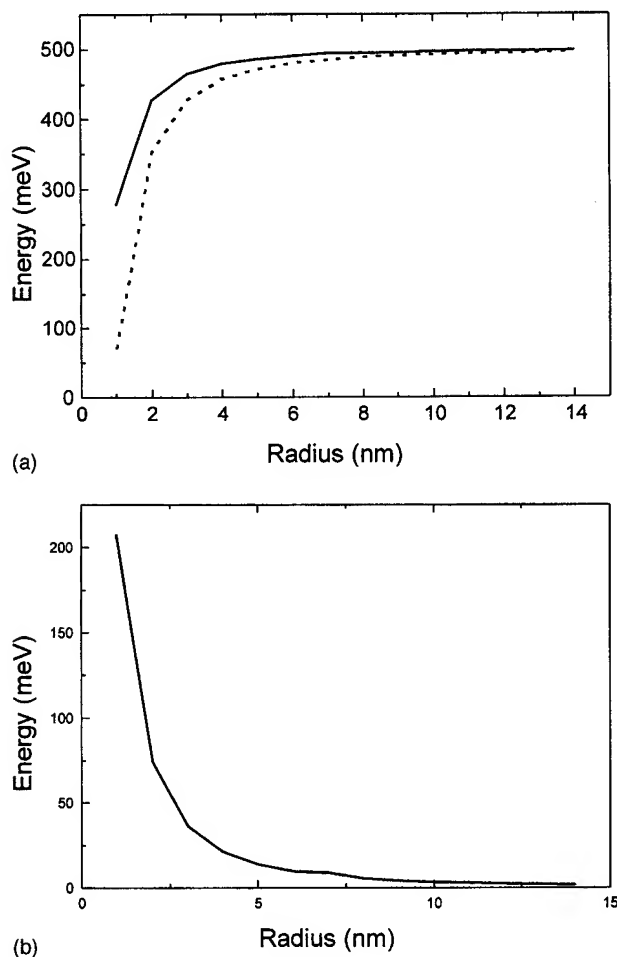


FIG. 2. Calculated orbital energies of the square spherical artificial atom well ($V_0 = 0.5$ eV) as a function of artificial atom radius. (a) The solid curve and the dot curve represent $1s$, $2p$ distinct energy level, respectively. (b) The differential value of the $1s \rightarrow 2p$ distinct energy level from (a).

levels. For example, as $R_0 = 5$ nm and $V_0 = 5$ eV, $\Delta E(1s - 2p) = 0.014$, $\Delta E(2p - 3d) = 0.018$; $\Delta E(5s - 6p) = 0.064$, and $\Delta E(6p - 7d) = 0.066$, where ΔE is in eV, while for a normal atom, the separation of energy levels in highly excited states is small and becomes continuous as $n \rightarrow \infty$. It shall be pointed out that the separation of neighboring energy levels in the same set is on the order of 10^2 meV for $R_0 = 5$ nm, 10 meV for $R_0 = 10$ nm, and 0.1 meV for $R_0 = 100$ nm. Quantum effects progressively become weak with increasing R_0 [Fig. 2(a)]. The differential value $1s \rightarrow 2p$ is decreased by increasing R_0 [Fig. 2(b)]. Our separations are close to 2 meV given in Ref. 12 ($R_0 = 100$ nm).

B. Circular cylindrical artificial atom well

The parameters used in the calculations are $a = 2.75$ nm and $l = 4$ nm. Energy levels calculated with different V_0 values are given in Table II. The level structure is more complex than that of the spherical artificial atom because every level is composed of two parts electronic motion along the Z direction in a one dimension potential well and electronic

TABLE II. The calculated energy levels of the artificial circular cylindrical atom (in meV).

V_0	K	$-E_{xy}$			$-E_z$
		$m=0$	$m=1$	$m=2$	
0.01	1	1.9137			4.6446
0.10	1	80.8139	52.373	17.3116	86.4311
	2	900648			48.0103
	3				0.7459
	1	464.1109	405.7865	348.9939	481.8801
	2	325.6018	88.3707	207.2912	427.8638
	3	191.5944			339.1568
	4				218.6486
	5				74.8974
	1	966.2114	914.1129	845.4926	980.5031
	2	821.2394	710.2616	581.4373	922.1534
	3	556.7039	365.7624	306.4541	825.4101
	4	286.0433		78.4369	691.1884
	5	65.6665			521.12235
	6				319.1893
	7				96.3246

motion on the XY plane in a two dimension potential well. There are no quantum numbers to indicate an electronic state; but the level structure has properties similar to that of the spherical artificial atom well. For example, when $V_0 \geq 0.2$ meV, the bound states appear; the number of the bound states is finite, etc.

IV. SUMMARY

In summary, we have dealt with the energy level structure of an artificial atom through the calculation of quantum mechanics and have obtained fundamental understanding on the energy level structure for this novel atom. Our calculated energy level sequence with a square spherical artificial atom well is exactly consistent with the theoretical sequence of Ref. 10, and energy level separations are in rough agreement with the experimental result of Ref. 12. The calculation of the circular cylindrical artificial atom well indicated that the distinction of change density appeared to be a periodic fluctuation along the Z direction or radius direction of the XY plane. In addition, this fluctuation period will change with electron energy state.

¹H. Chang, R. Grundbacher, D. Jovanovic, J. P. Leburton, and I. Adeslada, J. Appl. Phys. **76**, 3029 (1994).

²A. P. Alivisatos, Science **271**, 933 (1996).

³C. B. Murray, C. R. Kagan, and M. G. Bawendi, Science **270**, 1335 (1995).

⁴M. A. Kastner, Phys. Today **24**, (1993).

⁵M. F. Crommie, C. P. Lutz, and D. M. Eigler, Science **262**, 218 (1993).

⁶H. Masuda and K. Fukuda, Science **268**, 1466 (1995).

⁷D. H. Pearson and R. J. Tonucci, Science **270**, 68 (1995).

⁸H. Wu, D. W. L. Sprung, J. Martorell, and S. Klarsfeld, Phys. Rev. B **44**, 6351 (1991).

⁹C. M. Hsiao, W. N. Mei, and D. S. Chuu, Solid State Commun. **81**, 807 (1992).

¹⁰T. Inoshita, S. Ohnishi, and A. Oshiyama, Phys. Rev. Lett. **57**, 2560 (1986).

¹¹D. P. Divincenzo, Science **270**, 255 (1995).

¹²T. Demel, D. Heitmann, P. Grambow, and K. Ploog, Phys. Rev. Lett. **64**, 788 (1990).

Preparation and structural characterization of nanostructured CoAg granular films

H. Sang,^{a)} G. Ni, J. Lu, H. Chen, J. R. Zhang, and Y. W. Du

Institute for Solid State Physics and National Laboratory of Solid State Microstructures, Nanjing University, Nanjing 210093, China and Center for Advanced Studies in Science and Technology of Microstructures, Nanjing 210093, China

(Received 10 March 1997; accepted 17 March 1997)

The microstructures of as-deposited and annealed CoAg granular films fabricated by ion-beam cosputtering technique were characterized using x-ray diffraction, transmission electron microscopy (TEM), and ferromagnetic resonance (FMR) methods. For the $\text{Co}_{22}\text{Ag}_{78}$ sample, the changes in size and shape of the nanoscale cobalt granules embedded in the film closely depend on the annealing temperature (T_A). Real time observation *in situ* of TEM together with FMR spectra indicate that the size and shape of the cobalt granules evolve primarily along the film plane during the annealing process. Also, the FMR results indicate that the cobalt granules are still single domain particles embedded in the film as the sample is annealed up to 700 K. © 1997 American Vacuum Society. [S0734-211X(97)02104-5]

I. INTRODUCTION

Magnetic granular films are a kind of composite material in which nanoscale magnetic granules are usually embedded in an immiscible metallic or an insulator matrix at equilibrium. They display unusual magnetic and transport properties due to their unique microstructures. The advantage of the materials is that their properties can be artificially tailored in terms of fabricating conditions and post-fabrication treatment or annealing.¹

In the past few years, the giant magnetoresistance (GMR) effect in the magnetic granular films has been widely studied for both fundamental research and potential applications. The results show that the GMR effect in magnetic granular films is strongly related to the concentration of the magnetic element, the fabrication procedures, and the size of the magnetic granules. Among the many magnetic granular films, the CoAg system displays a particularly remarkable GMR effect.² Compared to layer-structured magnetic films (sandwich or multilayers), magnetic granular films have a more complicated microstructure. In the dilute magnetic system particularly, phase separation can lead to single domain nanoscale magnetic granules embedded in a matrix during fabrication using sputtering techniques, and the system displays superparamagnetism behavior above the blocking temperature.^{1,3} Experimentally, the optimal values of GMR often appear in the systems with lower concentration of magnetic elements (15–25 vol % or 20–30 at.%), and reach their maximum values after proper annealing.^{4–8} One believes that GMR originates from spin-dependent scattering of transport electrons, and the scattering from the surfaces of the magnetic granules is dominated.² Therefore, the microstructures of magnetic granular films play an important role for GMR, and intensive investigation of microstructures in detail is needed to understand the mechanism of GMR in these systems.

In this article, we mainly present the results of the studies of the fabrication and structural characterization of CoAg granular films. The dependence of GMR on the microstructures is also discussed.

II. EXPERIMENT

A series of CoAg granular films was prepared by the ion-beam cosputtering technique at a substrate temperature (T_S) of 300 K. The high-vacuum system is equipped with a Kaufman ion source⁹ in which argon (99.99%) pressure is maintained at about 0.2 mTorr. After being ionized in a high voltage chamber, the beam of argon ions with energy of 1.2 keV is drawn by an acceleration grid and projected directly at the water-cooled target at an incident angle of 45°. The target is in a mosaic style, consisting of pure silver and pure cobalt metals. The small square cobalt (or silver) pieces are glued to the silver (or cobalt) plate. Because of the different sputtering ratio of both cobalt and silver, the small square pieces glued to the plate can be arranged by both the number and the different patterns to change the cobalt concentration x . The sputtered materials were deposited onto substrates fixed in a water-cooled substrate holder installed in the specularly reflected direction of the incident ion beam. The distances from the target to both ion source and substrate holder are about 15 cm. X-ray diffraction (XRD), transmission electron microscopy (TEM), and ferromagnetic resonance (FMR) were employed to investigate the microstructures of the samples. The films were deposited onto common glass substrates with the thickness of the order of 400 nm for XRD structural analysis and for FMR spectrum measurements, and onto a copper grid which was covered previously with a pyroxylin colloid layer in thickness of a few tens of nanometers to support the film for TEM observation. The thickness of the TEM samples was in the range of 50–100 nm in order to be directly examined without a thinning procedure. The exact cobalt concentration, x , was determined using an electron microprobe.

^{a)}Corresponding author; Electronic mail: haisang@nju.edu.cn

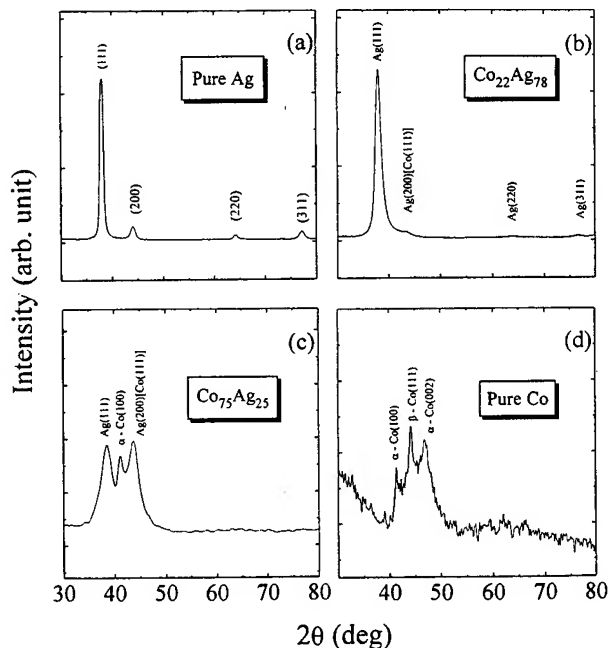


FIG. 1. XRD patterns for as-deposited CoAg samples ($T_s=300$ K), indicating concentration-dependent effect.

III. RESULTS AND DISCUSSION

Figure 1 shows a set of XRD patterns for as-deposited CoAg granular film samples prepared at $T_s=300$ K with different cobalt concentrations, showing the cobalt concentration-dependent effect. The XRD pattern of a pure Ag sample clearly shows a fcc polycrystalline structure with (111) texture, accompanying the (111), (200), (220), and (311) peaks. The diffraction peaks are sharp with narrow widths, indicating the existence of large size Ag grains in the sample [Fig. 1(a)]. After adding cobalt, the XRD patterns apparently change. At $x=22$ at. %, the Ag(200) peak changes to a shoulder, and the intensity of Ag(220), (311) peaks decrease [Fig. 1(b)]. No crystalline evidence for either fcc or hcp cobalt appears due to the overlap of the Ag(200) and Co(111) peaks, even though there is strong tendency for phase separation in CoAg samples. However, the broadening of the Ag(200) peak indicates the appearance of small size cobalt granules.¹⁰ When x reaches 75 at. %, the hcp cobalt structure [α -Co(100)] appears in the XRD pattern, and the intensity of the Ag(111) peak drastically decreases, whereas that of a narrow Ag(200)[Co(111)] peak increases somewhat. These results suggest that cobalt granules grow larger into islands and connect to form an extended network [Fig. 1(c)]. Even in pure cobalt single layer film, fcc, and hcp structures coexist [Fig. 1(d)].

We have also observed *in situ* during the annealing, the changes in the size and shape of the cobalt granules embedded in a silver matrix for the $\text{Co}_{22}\text{Ag}_{78}$ sample using an ultrahigh vacuum (UHV) TEM equipped with a real-time video recording system. This sample was chosen for direct observation because we previously obtained the largest room temperature GMR in this sample after a proper annealing treatment.¹¹ In this procedure, the $\text{Co}_{22}\text{Ag}_{78}$ sample depos-

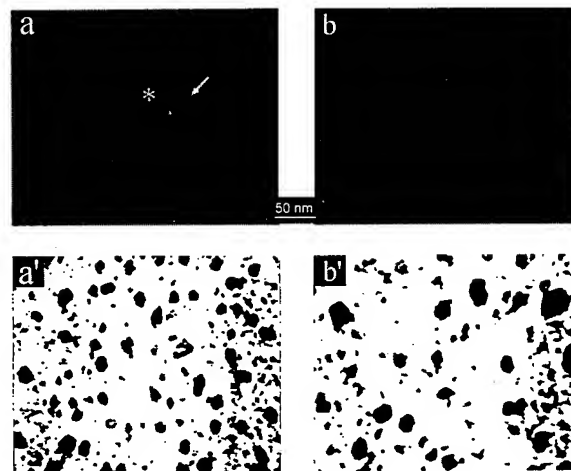


FIG. 2. A set of TEM bright field micrographs of the selected area: (a) and (b) corresponded to an as-deposited and a 10 min annealed $\text{Co}_{22}\text{Ag}_{78}$ sample, respectively. (a') and (b') are corresponding high contrast pictures after drawing process using the computer at the same contrast threshold.

ited on the copper grid was fixed on a thermal holder in which the temperature was controlled by a temperature controller. Annealing temperatures were 400, 500, 600, and 700 K, respectively, with a ramp rate of 20 K/min. Figure 2 shows the TEM bright field micrographs of an as-deposited sample [Fig. 2(a)] and the sample after being annealed up to 700 K [Fig. 2(b)]. The area was selected for observation because in this area one of the granules had a special shape that could be used as a reference point to easily observe *in situ* the changes in the size and shape of other nearby granules during annealing. Annealing time for each temperature was 10 min. These snapshots were selected from a video tape recorded over a period of more than 50 min at various annealing temperatures. Figures 2(a)' and 2(b)' correspond to Figs. 2(a) and 2(b), respectively, showing high contrast pictures drawn by computer at the same contrast threshold. They clearly show the changes of the size and shape of the cobalt granules with annealing. The average size of the granules, estimating from TEM results, increases from about 10 nm in the as-deposited state to about 15 nm at $T_A=700$ K, and the changes of the size and shape of the cobalt granules are more dependent on annealing temperature T_A than on annealing time.¹²

From the TEM images, one can see that not only the size but also the shape of the cobalt granules change. The TEM pictures, however, only provide two-dimensional microscopic images of the cobalt granules. In the sample with low cobalt concentration, shape-induced anisotropy is dominant. According to the Kittel relation for FMR:¹³

$$\left(\frac{\omega}{\gamma}\right)^2 = [H + 4\pi(N_x - N_z)M_{\text{eff}}][H + 4\pi(N_y - N_z)M_{\text{eff}}], \quad (1)$$

and

$$\left(\frac{\omega}{\gamma}\right)^2 = H_{\parallel}(H_{\parallel} + 4\pi M_{\text{eff}}) \left(\theta_H = \frac{\pi}{2}\right), \quad (2)$$

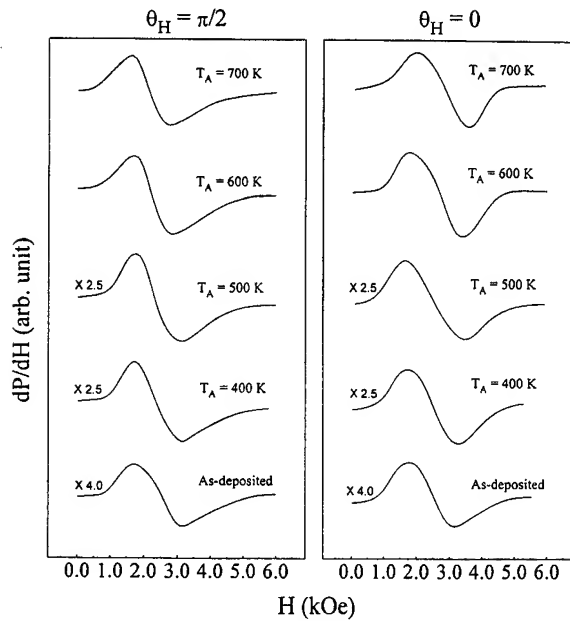


FIG. 3. The dependence of FMR spectra of the $\text{Co}_{22}\text{Ag}_{78}$ sample on annealing temperatures. $\theta_H = \pi/2$ and $\theta_H = 0$ expresses the applied field parallel and perpendicular to sample planes, respectively.

$$\frac{\omega}{\gamma} = H_{\perp} + 4\pi M_{\text{eff}} \quad (\theta_H = 0), \quad (3)$$

where θ_H is the angle between the applied field H and the film normal and M_{eff} is effective magnetization. Since the change of demagnetizing factor can reflect the variation of cobalt granule shapes, FMR results can provide three-dimensional information of the small ferromagnetic granules. The FMR spectra were measured at room temperature. In the experiment, the microwave field was applied parallel to the plane of the samples with a frequency of 9.8 GHz. Figure 3 shows two sets of FMR spectra for $\text{Co}_{22}\text{Ag}_{78}$ samples annealed at different temperatures. These sets correspond to data collected with the applied field parallel (H_{\parallel} , $\theta_H = \pi/2$) and perpendicular (H_{\perp} , $\theta_H = 0$) to the planes of the samples, respectively. In the as-deposited state, $H_{r\parallel} \approx H_{r\perp}$. The difference of resonance field, $\Delta H_r = H_{r\perp} - H_{r\parallel}$, increases monotonically as T_A rises. The variation of ΔH_r indicates that the cobalt granules in as-deposited sample were roughly spherical in shape in terms of Eq. (1). As T_A rises, $H_{r\perp}$ moves toward high field and $H_{r\parallel}$ towards low field, which means that the demagnetizing field increases perpendicular to the plane of sample and decreases parallel to the plane. It suggests that the shapes of the cobalt granules evolve mainly in-plane from roughly spherical to pancakelike during annealing.

From Fig. 3, one can also see that the linewidth of the resonance peak (peak-peak) gradually change as T_A increases. By intensive investigation of FMR spectra of powder-pattern ferromagnetic particles, Griscom *et al.*^{14,15} found that the linewidth can be described as

$$(\Delta H)_{\text{SD}} = \frac{10K}{3M} \quad (4)$$

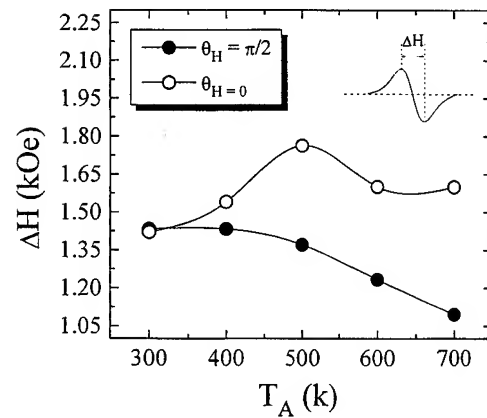


FIG. 4. Linewidths of FMR spectra shown in Fig. 3 as functions of T_A . Lines were drawn for a guide to the eyes.

for single-domain particles, and as

$$(\Delta H)_{\text{MD}} = 0.7 \frac{4\pi M}{3} \quad (5)$$

for multidomain particles, where K and M are the magnetic anisotropy energy and magnetization of a ferromagnetic particle, respectively. In principle, dilute magnetic granular films (e.g., the $\text{Co}_{22}\text{Ag}_{78}$ sample) may exhibit the character of ultrafine particles, and the average size of the cobalt granules is only on the order of 15 nm in the $\text{Co}_{22}\text{Ag}_{78}$ sample after annealing up to 700 K. For pure fcc structural cobalt, $K/M \sim 0.6$ kOe and $4\pi M \sim 17$ kOe, and yield the theoretical linewidths of 2.0 kOe for single-domain and 4.0 kOe for multidomain particles according to Eqs. (3) and (4). Figure 4 presents the linewidths obtained from the FMR spectra in Fig. 3 as functions of T_A . One can see that the linewidths monotonically decrease for $\theta_H = \pi/2$ and slightly rise and fall for $\theta_H = 0$ as T_A increases. Experimental linewidths shown in Fig. 4 are all smaller than the theoretical linewidths for the single-domain linewidth, with a maximum value of about 1.8 kOe. Therefore, this indicates that there only exists single-domain cobalt granules in the $\text{Co}_{22}\text{Ag}_{78}$ sample after annealing up to 700 K for 10 min.

The evolution of the cobalt granules in the sample under annealing treatment strongly influences the GMR. The optimal value of GMR appears at $T_A = 500$ K in the $\text{Co}_{22}\text{Ag}_{78}$ sample.¹¹ The strong dependence of GMR on annealing temperature indicates that the microstructures, and both shape and size of the cobalt granules, play an important role. In the region of $300 \leq T_A < 500$ K, the change of the cobalt granule size is not obvious.¹² The MR increment is mainly due to the change of roughness at the interface between the cobalt granules and the silver matrix. After annealing, the interface roughness is improved. As is well known, certain roughness favors GMR.¹⁰

In conclusion, we have reported some results of the microstructures in as-deposited and 10 min annealed CoAg granular films fabricated by ion-beam cosputtering technique. For the $\text{Co}_{22}\text{Ag}_{78}$ sample, the changes in size and shape of the cobalt granules in the film are closely related to

annealing temperature. The results of TEM real-time observation and FMR indicate that the size and shape of the cobalt granules evolve primarily in the film plane during the annealing procedure. The FMR results also show that there only exist single domain cobalt granules in the sample after annealing up to 700 K.

ACKNOWLEDGMENTS

The authors would like to thank X. N. Zhao, J. M. Hong, S. Y. Han, and Y. X. Sui for their assistance in TEM observations and FMR spectrum measurements. One of the authors (H. Sang) would like to thank the support of Ke-li Foundation and thank S. P. Wong, N. Ke, and W. Y. Cheung for their valuable discussions. This work was supported in part by Grant No. 85-6 NMS, Grant NSFC, and Grant JSNSF.

¹C. L. Chien, in *Science and Technology of Nanostructured Magnetic Material*, edited by G. C. Hadjipanayis and G. A. Prinz (Plenum, New York, 1991), p. 477.

²Jian Qing Wang and Gang Xiao, *Phys. Rev. B* **49**, 3982 (1994).

³J. R. Childress and C. L. Chien, *J. Appl. Phys.* **70**, 5885 (1991).

⁴J. Q. Xiao, J. S. Jiang, and C. L. Chien, *Phys. Rev. Lett.* **68**, 3749 (1992).

⁵A. E. Berkowitz, J. R. Mitchell, M. J. Carey, A. D. Young, S. Zhang, F. E. Spada, F. T. Parker, A. Hutten, and G. Thomas, *Phys. Rev. Lett.* **68**, 3745 (1992).

⁶J. A. Barnard, A. Wakis, M. Tan, E. Haftek, M. R. Parker, and M. L. Watson, *J. Magn. Magn. Mater.* **114**, L230 (1992); J. A. Barnard, S. Hossain, M. R. Parker, A. Wakis, and M. L. Watson, *J. Appl. Phys.* **73**, 6372 (1993).

⁷J. Q. Xiao, J. S. Jiang, and C. L. Chien, *Phys. Rev. B* **46**, 9266 (1992).

⁸Gang Xiao, Jian Qing Wang, and Peng Xiong, *Appl. Phys. Lett.* **62**, 420 (1993).

⁹H. R. Kaufman, *Electron Phys.* **36**, 265 (1974).

¹⁰H. Sang, N. Xu, J. H. Du, G. Ni, S. Y. Zhang, and Y. W. Du, *Phys. Rev. B* **53**, 15023 (1996).

¹¹H. Sang, Z. S. Jiang, G. Guo, J. T. Ji, S. Y. Zhang, and Y. W. Du, *J. Magn. Magn. Mater.* **140-144**, 589 (1995).

¹²H. Sang, S. Y. Zhang, H. Chen, G. Ni, J. M. Hong, X. N. Zhao, Z. S. Jiang, and Y. W. Du, *Appl. Phys. Lett.* **67**, 2017 (1995).

¹³C. Kittel, *Phys. Rev.* **73**, 155 (1948).

¹⁴D. L. Griscom, C. L. Marquardt, E. J. Friebele, and D. J. Dunlop, *Earth Planet. Sci. Lett.* **24**, 78 (1974).

¹⁵M. Rubenstein, B. N. Das, N. C. Koon, D. B. Chrissey, and J. Horowitz, *Phys. Rev. B* **50**, 184 (1994).

Fabrication, characterization, and the photoelectric conversion of the nanostructured TiO₂ electrode

Huihua Deng^{a)} and Zuhong Lu

National Laboratory of Molecular and Biomolecular Electronics, Southeast University, Nanjing 210096, China and Center for Advanced Studies in Science and Technology of Microstructures, Nanjing 210093, China

Haifang Mao and Huijun Xu

Institute of Photographic Chemistry, Academic Sinica, Beijing 100101, China

(Received 10 March 1997; accepted 10 March 1997)

High-surface-area TiO₂ film has been deposited on an ITO conducting glass substrate from colloidal suspension. Electronic contact between particles is produced by sintering at 450 °C. Atomic force microscope and scanning electron microscope analyses show that the TiO₂ electrode is made up of nanometer-sized and interconnected TiO₂ particles and pores. The specific surface area of the TiO₂ electrode is determined by the size of particles and the thickness of the film and can be measured to be 150 by the absorption spectrum study of the TiO₂ electrode sensitized with zinc tetrasulfonated phthalocyanine (ZnTsPc). A liquid junction cell based on the ZnTsPc/nanostructured TiO₂ electrode has been fabricated and harvests 20% of the incident solar energy flux. Under monochromatic illumination at $\lambda=700$ nm, the incident-photon-to-current conversion efficiency is about 2.1% and the quantum efficiency, considering the actual absorption of the incident light, is 7%. The conversion efficiency reaches 1.7% under an illumination of 35 mW cm⁻², which presents one of the highest values reported for phthalocyanine photovoltaic devices. © 1997 American Vacuum Society. [S0734-211X(97)01904-5]

I. INTRODUCTION

Organic photovoltaic solar cells have been intensively investigated. The major advantages of organic cells, in which an organic dye layer is used to extend the absorbance of a solar cell into the visible region, are the full use of solar energy and the ease of fabrication and low cost. These devices conventionally have a sandwich structure in which an organic layer is interposed between a metal and a semiconductor (or metal) electrode.¹⁻⁴ So far, however, the energy conversion efficiency is poor at about 1%. Whether in the form of a photoelectrochemical cell or a solid cell, the low efficiency in organic photovoltaic cells is due to the low efficiency of charge carrier separation and transport if the dye layer is thick^{3,4} or due to the low absorbance if the dye layer is thin.⁵ The conventional approach to improving the energy conversion efficiency is to increase the conductivity of the dye by doping with an appropriate sensitizer such as iodine and AgI.^{3,6} Recently, a Swiss research group has developed several concepts in order to overcome these two limitations. The surface area of the transparent electrode was increased so that a large number of dye molecules can be adsorbed directly into the electrode surface and simultaneously be in direct contact with the redox electrolyte, which results in efficient separation and transport of the photogenerated charge carriers.⁷ If a sufficiently large surface area of the electrode can be provided, even a monolayer of dye molecules adsorbed on the electrode could absorb most of the incident photons.⁸ In our laboratory, the TiO₂ microporous

electrode with a specific surface area of 50 had been fabricated.⁹ By adding hydroquinone to the liquid electrolyte that contains 0.1 M Na₂SO₄, the quantum efficiency in photoelectric conversion of the microporous TiO₂ electrode, sensitized with zinc phthalocyanine substituted with a carboxyl group, reaches about 20% at 690 nm.¹⁰ However, the energy conversion efficiency of the liquid junction cell based on the nanostructured TiO₂ electrode sensitized with phthalocyanine is low. One reason is the fact that the absorbance of the phthalocyanine dimer does not convert into a photocurrent.¹¹ It is necessary to further investigate other factors giving rise to low efficiency. In this article, the nanostructured TiO₂ electrode with a large specific surface area of 150 is fabricated. Zinc tetrasulfonated phthalocyanine (ZnTsPc) was used to characterize the specific surface area of the electrode since ZnTsPc self-assembled on the nanostructured TiO₂ electrode via the negatively charged sulfonate (SO₃⁻) and its photoelectric property is best for MTsPc (M=Zn, In, Ga, Co, TiO, H₂).¹¹ In addition, the photoelectric conversion of the liquid junction cell based on the nanostructured TiO₂ electrode sensitized with ZnTsPc is also presented, from which it is found that the formation of the liquid junction between the TiO₂ electrode and electrolyte is important in improving energy conversion.

II. EXPERIMENTAL DETAILS

A. Fabrication of the conducting TiO₂ thin layer

TiO₂ colloidal suspension was prepared by hydrolysis of tetrabutyl titanate [(C₄H₉O)₄Ti] by a procedure similar to that described in Ref. 7. 2% poly(vinyl alcohol) (PVA) was

^{a)}Author to whom correspondence should be addressed; Electronic mail: deng@seu.edu.cn

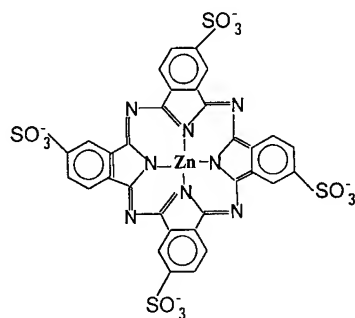


FIG. 1. The molecular structure of zinc tetrasulfonated phthalocyanine (ZnTsPc).

added to the suspension, which was then concentrated by vacuum rotation evaporation. Before TiO₂ suspension was spread, a Triton monolayer, formed on a double-distilled water surface on a Langmuir trough, was deposited on freshly cleaned indium tin oxide [ITO; sheet resistance of 50 Ω/□ and 90% transmission in the visible region] conducting glass substrates through the vertical dipping method. The surface pressure was kept at 15 mN m⁻¹ and the dipping speed was about 5 mm min⁻¹. The concentrated suspension (TiO₂ content 10% by weight) in another Langmuir trough was spread on the above ITO conducting glass substrate through the horizontal dipping method. The dipping speed was about 2.5 mm min⁻¹. The film was heated in air in a furnace in which the temperature was increased gradually to 450 °C and then kept at 450 °C for 30 min. The TiO₂ film, obtained by finally annealing at 550 °C under nitrogen gas flow for another half an hour, possesses the resistivity of 500 Ω/□ and a thickness of 100 nm. The function of the conducting TiO₂ thin layer lies in the formation of the liquid junction in the interface between the TiO₂ electrode and electrolyte.

B. Fabrication of the nanostructured TiO₂ electrode

The viscous suspension the a TiO₂ content of 20%–30% by weight and a PVA content of 10% by weight was spin-coated on the above conducting glass substrate. The nanostructured TiO₂ electrode was obtained by heating in air for 30 min at 450 °C. Before dye coatings, the nanostructured TiO₂ electrodes were soaked in a HCl solution (pH=2) and naturally dried.

The zinc tetrasulfonated phthalocyanine (ZnTsPc) was synthesized according to the method described in Ref. 12 and the molecular structure is shown in Fig. 1. ZnTsPc self-assembles on the nanostructured TiO₂ electrode via the electrostatic interaction between the negatively charged sulfonate (SO₃⁻) and the positively charged surface of TiO₂ nanoparticles, which is demonstrated by the experimental results that the free-metal tetrasulfonated phthalocyanine also self-assembles the TiO₂ electrode acidified by 0.01 M HCl solution and ZnTsPc does not adsorb on the TiO₂ electrode soaked in 0.001 M NaOH solution.¹¹ The working electrode was obtained by plunging the TiO₂ electrode into a

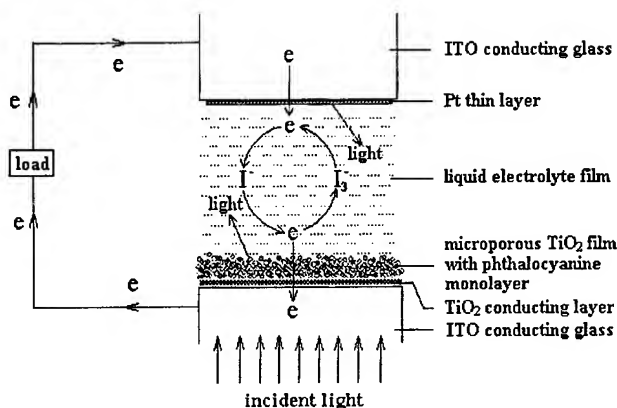


FIG. 2. The structure of the ZnTsPc/TiO₂/electrolyte liquid junction cell based on the nanostructured TiO₂ electrode.

5 × 10⁻⁴ M solution of ZnTsPc in dimethyl sulfoxide (DMSO) for 1 h. The electrode became green after sensitization with ZnTsPc.

C. Fabrication of the liquid junction cell based on the nanostructured TiO₂ electrode

To form the final cell, a drop of liquid electrolyte, which contains 0.1 M KI and 0.05 M iodine in 0.001 M HClO₄ solution, is placed on the film and a thin layer of a redox electrolyte percolates into the pores of the nanostructured TiO₂ electrode through capillary forces. The liquid junction between the electrode and electrolyte is generated because a thin layer of the closely packed TiO₂ film prevents the direct contact of the liquid electrolyte with an ITO conducting layer. A counter electrode of the ITO conductive glass, on which a thin layer of platinum was coated by physical vacuum deposition at 10⁻⁷ Torr, is placed on top and the sandwich is illuminated through the TiO₂ support as shown in Fig. 2.

D. Measurements

The morphology of the TiO₂ electrode was examined by scanning electron microscope [(SEM); JEOL, JSM-6300] and atomic force microscope [(AFM); DI, Co., Nanoscope III]. The size of the TiO₂ colloidal particles was observed by transmission electron microscope [(TEM); Hitachi, JEM-2000EX]. The absorption spectra of the ZnTsPc/TiO₂ electrode were recorded with a Shimadzu UV-2201 UV-visible spectrophotometer. The photocurrent and photovoltage were measured with a potentiostat model CMBP-1. Monochromatic illumination was obtained using a 500 W xenon arc lamp in combination with a grating monochromator WPG3D model. The intensity was calibrated using a model FP-3 radiometer-photometer.

III. RESULTS AND DISCUSSION

A. Morphology and specific surface area of the nanostructured TiO₂ electrode

The conducting TiO₂ thin layer in Fig. 3(a) consists of the close cubic packing of TiO₂ nanoparticles. The nanostruc-

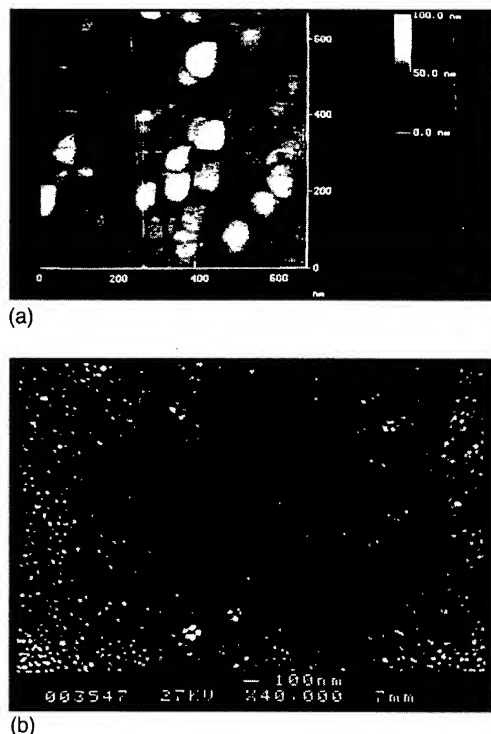


FIG. 3. (a) Morphology of the conducting TiO₂ thin layer by AFM (650 nm×650 nm) and (b) surface morphology of the nanostructured TiO₂ electrode by SEM.

tured TiO₂ electrode in Fig. 3(b) is found to have a thickness of 10 μ m and is composed of interconnected and nanometer-sized TiO₂ particles and micropores. It can also be seen that the TiO₂ particles in the film are uniform with an average diameter of 55 nm, which is much larger than the size of particles in colloidal solution (3–4 nm as determined by TEM). The size of the particles and pores making up the film is controlled by the size of the particles in the colloidal suspension.¹⁰ The internal surface area of the film is determined on the size of particles and the thickness of the film.^{7,8} These parameters were optimized to obtain efficient light harvesting while maintaining a pore size large enough to allow the dye molecules to penetrate into the internal surface of the nanostructured TiO₂ electrode and allowing the redox electrolyte to diffuse easily. This optimum successfully realizes the direct adsorption of dye molecules on the surface of the electrode and direct contact between electrolyte and dye molecules. A cubic close packing of 55-nm-sized spheres to a 10- μ m-thick layer is expected to produce a 550-fold increase in surface area. The specific surface area, which is defined as the ratio of the total surface area of TiO₂ nanoparticles to the geometrical surface area of the TiO₂ film in the electrode, is expected to reach 550.

The absorption spectra of the TiO₂ electrode sensitized without (curve a) and with (curve b) ZnTsPc have been measured and are shown in Fig. 4. The bare TiO₂ films are transparent and colorless, displaying the fundamental absorption onset of anatase at 390 nm (corresponding to a band gap energy of 3.2 eV) in the UV region. The ZnTsPc/TiO₂ elec-

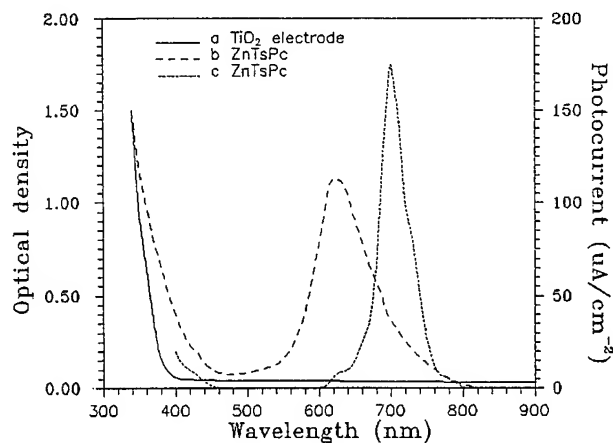


FIG. 4. Absorption spectra of (a) the bare TiO₂ electrode and (b) the nanostructured TiO₂ electrode sensitized with ZnTsPc, using an ITO conducting glass as a reference; and (c) photocurrent action spectrum of the liquid junction cell based on the ZnTsPc/TiO₂ electrode.

trode reveals the characteristic absorption between 580 and 750 nm in the visible region, indicating that the sensitization of the TiO₂ electrode with ZnTsPc dye molecules could extend the absorbance of the electrode into the visible region. Integration of the spectral overlap between a light source and this absorption band shows that 20% of the incident light energy flux is harvested by the ZnTsPc-sensitized TiO₂ film.

In general, the peak at 620 nm can be assigned as the dimeric ZnTsPc *Q* band and the absorbance at 700 nm can be attributed to the monomeric ZnTsPc *Q* band.¹³ At λ =620 nm, the optical density of the ZnTsPc/TiO₂ electrode and the extinction coefficient¹⁴ of the ZnTsPc dimer are about 1.26 and $\sim 10^8$ cm² mol⁻¹, respectively. At λ =700 nm, the optical density of the ZnTsPc/TiO₂ electrode and the extinction coefficient¹¹ of the ZnTsPc monomer are 0.34 and 3.3×10^9 cm² mol⁻¹, respectively. The surface concentrations of the ZnTsPc dimer and monomer adsorbed on the electrode are 1.26×10^{-8} and 1.11×10^{-10} mol cm⁻², respectively. Therefore, ZnTsPc adsorbed on the nanostructured TiO₂ electrode mainly exists as a dimer.

The specific surface area of the nanostructured TiO₂ electrode is characterized by the absorbance of the ZnTsPc dimer and can be derived from a method similar to that described in Ref. 8 by assuming complete monolayer coverage on the TiO₂ electrode and a 2 nm² area for each molecule. The specific surface area for the ZnTsPc/TiO₂ nanostructured electrode is 150, which is smaller than the predicted values of 550. The difference is attributed to necking between particles. In addition, a capillary force prevents the access of large ZnTsPc aggregates to very small pores, reducing the apparent surface area.

B. Photoelectric conversion in the liquid junction cell based on the ZnTsPc/TiO₂ nanostructured electrode

Curve c in Fig. 4 shows the measured short-circuit photocurrent of the ZnTsPc/TiO₂ electrode as a function of wavelength. The spectrum has been corrected for the absorp-

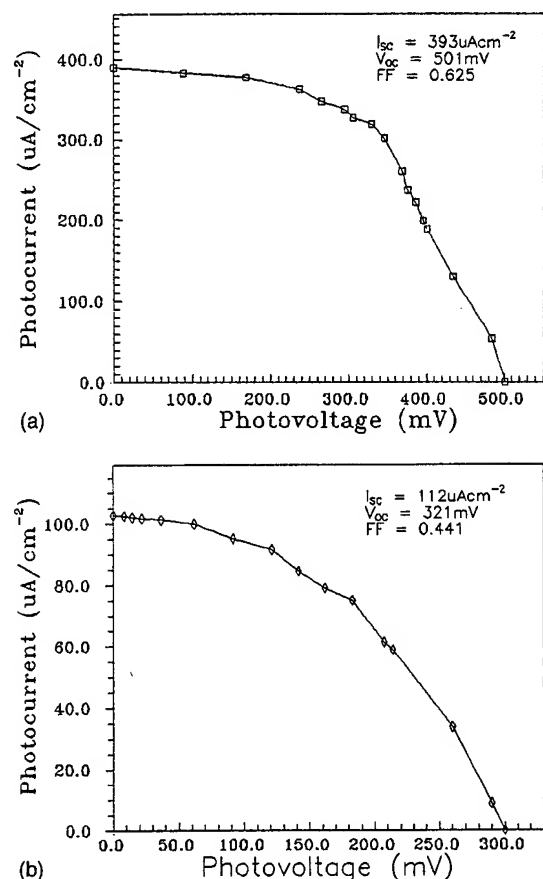


FIG. 5. The photovoltage-current characteristics of the liquid junction cell based on the ZnTsPc/TiO₂ nanostructured electrode under an illumination of 35 mW cm⁻² with (a) and without (b) the pre-deposition of the conducting TiO₂ thin layer on an ITO conducting glass

tion of incident light by the conducting ITO glass substrate. The incident-photon-to-current conversion efficiency (IPCE) of 2.1% at $\lambda=700$ nm is calculated from the formula⁷ $IPCE = 1243 \times I_{SC} \times 100\% / (\lambda \times P_{in})$, in which I_{SC} and P_{in} are the short-circuit photocurrent density and the incident light power at a monochromatic wavelength, respectively. Considering the actual part (R_{ab}) of the incident light absorbed by the dye molecules, the quantum efficiency (η) reaches 7%, which can be derived from the formula¹¹ $\eta = IPCE / R_{ab}$ at a monochromatic wavelength.

Figure 5 shows the photovoltage-current characteristics of the liquid junction cell based on the ZnTsPc/TiO₂ nanostructured electrode under an illumination of 35 mW cm⁻². The conversion efficiencies without and with the conducting TiO₂ thin layer are 0.25% and 1.7%, and the fill factors ((FF); maximum output power of the cell ÷ [short-circuit current (I_{SC}) × open-circuit voltage (V_{OC})] are 0.441 and 0.625, respectively. A marked improvement in the photocurrent (I_{SC}) and energy conversion is achieved through the pre-deposition of the conducting TiO₂ thin layer on an ITO conducting glass. The conducting TiO₂ thin layer in Fig. 3(a) consists of the close cubic packing of TiO₂ particles and the liquid junction in the interface between the TiO₂ electrode and electrolyte is generated. The band bending resulting

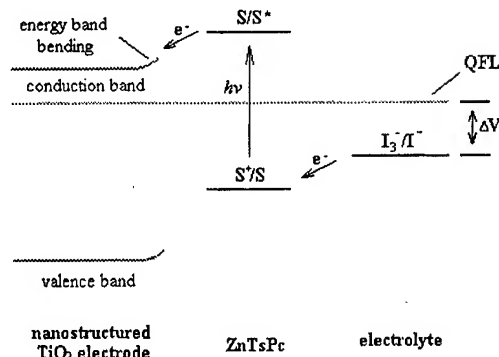


FIG. 6. Electron injection from the excited sensitizer S^* (ZnTsPc^{*}) into the conduction band of TiO₂ and electron transfer from liquid electrolyte (I_3^-/I^-) to the oxidized sensitizer S^+ (ZnTsPc⁺). QFL, quasi-Fermi level of TiO₂ under illumination. ΔV , the difference between the QFL and the electrochemical potential of the electrolyte that is equal to the Nernst potential of the redox couple (I_3^-/I^-); $h\nu$, the incident photon energy.

from the liquid junction, as shown in Fig. 6, aids electron injection from the excited state of ZnTsPc into the conduction band of the TiO₂ semiconductor. The band energies of TiO₂ and ZnTsPc are taken from Refs. 15 and 16. In the meantime, the conducting TiO₂ layer with the closely packed structure prevents electrons of the redox electrolyte (I_3^-/I^-) from directly being transported into the ITO conducting layer and keeps the potential difference (ΔV) constant. A redox electrolyte (I_3^-/I^-) is used to mediate charge transfer between the electrodes and to regenerate the sensitizer. ΔV in Fig. 6 corresponds to the difference between the quasi-Fermi level in TiO₂ under illumination and the Nernst potential of the electrolyte. Thus, ZnTsPc regenerates efficiently by electron transfer from a redox electrolyte in solution. In addition, the nanometer-sized TiO₂ particles in the nanostructured TiO₂ electrode allow efficient transport and rapid separation of the photogenerated charge carriers and minimize surface and bulk recombination.

The fill factor of the cell in Fig. 5(a) reaches 0.625, which is higher than that (<0.5) of conventional photovoltaic cells and the cell in Fig. 5(b). This indicates that loss mechanisms such as recombination, normally encountered in semiconductor photoconversion, have been minimized. The role of TiO₂ in the liquid junction cell based on the ZnTsPc/TiO₂ nanostructured electrode is merely to conduct the injected majority charge carriers (here it refers to electrons). There are no minority carriers involved in the photoconversion process. Surface and bulk recombination losses due to lattice defects, encountered in conventional photovoltaic cells, are not observable in such a device.

IV. SUMMARY

We have prepared the nanostructured TiO₂ electrode with a high surface area of 150, which is characterized by the dimeric absorbance of zinc tetrasulfonated phthalocyanine self-assembled on the TiO₂ electrode. A liquid junction cell

based on the ZnTsPc/TiO₂ nanostructured electrode has been fabricated and harvests 20% of the incident solar energy flux. Under monochromatic illumination at $\lambda=700$ nm, the incident-photon-to-current conversion efficiency is about 2.1% and the quantum efficiency, considering the actual absorption of the incident light, is 7%. The conversion efficiency reaches 1.7% under an illumination of 35 mW cm⁻², which presents one of the highest values reported for phthalocyanine photovoltaic devices. The predeposition of the conducting TiO₂ thin layer with the closely packed structure aids the formation of the liquid junction in the interface between the TiO₂ electrode and liquid electrolyte and results in the enhancement of photoelectric performances of the solar cell.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China.

- ¹A. K. Ghosh, D. L. Morel, T. Feng, R. S. Shaw, and C. A. Rowe, *J. Appl. Phys.* **45**, 230 (1974).
- ²F. R. Fan and L. R. Faulkner, *J. Chem. Phys.* **69**, 3334 (1978).
- ³G. D. Sharma, S. C. Mathur, and D. C. Dube, *J. Mater. Sci.* **26**, 6547 (1991).
- ⁴W. A. Nevin and G. A. Chamberlain, *J. Appl. Phys.* **69**, 4324 (1991).
- ⁵H. Gerischer, in *Photoelectrochemistry, Photocatalysis and Photoreaction*, edited by M. Schiavello (Reidel, Dordrecht, 1985).
- ⁶W. A. Nevin and G. A. Chamberlain, *J. Appl. Phys.* **68**, 5247 (1990).
- ⁷B. O'Regan, J. Moser, M. Anderson, and M. Gratzel, *J. Phys. Chem.* **94**, 8720 (1990).
- ⁸B. O'Regan and M. Gratzel, *Nature (London)* **253**, 737 (1991).
- ⁹Y. C. Shen, Z. H. Lu, and Y. Wei, *Chin. Sci. Bull.* **39**, 2238 (1994).
- ¹⁰Y. C. Shen, L. Wang, Z. H. Lu, Y. Wei, Q. F. Zhou, H. F. Mao, and H. J. Xu, *Thin Solid Films* **257**, 144 (1995).
- ¹¹H. H. Deng, H. F. Mao, B. J. Liang, Y. C. Shen, Z. H. Lu, and H. J. Xu, *J. Photochem. Photobiol. A* **99**, 71 (1996).
- ¹²H. Ali, R. Langlois, and J. R. Wanger, *Photochem. Photobiol.* **47**, 713 (1988).
- ¹³W. A. Nevin, W. Liu, and A. B. P. Lever, *Can. J. Chem.* **65**, 855 (1987).
- ¹⁴L. D. Rollman and R. T. Iwamoto, *J. Am. Chem. Soc.* **90**, 1455 (1968).
- ¹⁵F. F. Fan and A. J. Bard, *J. Am. Chem. Soc.* **101**, 6137 (1979).
- ¹⁶J. R. Darwent, I. McCubbin, and D. Phillip, *J. Chem. Soc. Faraday Trans.* **78**, 347 (1982).

Surface vibrational mode of ZnS nanoparticles

Jianfeng Xu,^{a)} Haitao Mao, and Yue Sun
Department of Physics, Henan University, Kaifeng 475001, China

Youwei Du

Laboratory of Solid State Microstructures, Nanjing University, Nanjing 210093, China

(Received 8 September 1996; accepted 5 March 1997)

Raman scattering and infrared absorption from ZnS nanoparticles were measured. A broad peak was observed in Raman and infrared absorption spectra which can be attributed to the surface vibrational mode of ZnS nanoparticles. © 1997 American Vacuum Society. [S0734-211X(97)02304-4]

I. INTRODUCTION

Nanoparticles exhibit various new properties that differ from the natural bulk materials due to surface and size effects.¹⁻⁴ It has been argued that many remarkable phenomena such as an enhancement of the specific heat at low temperatures⁵ and a much lower melting point than that of corresponding bulk material⁶ are related to the modulation of the phonon spectrum. Recently, the nature of vibrations of nanoparticles thus has received a great deal of attention from researchers both theoretically and experimentally.^{7,8} For nanoparticles, the surface to volume ratio is very large, and the number of atoms on the surface is comparable to or larger than that inside the particles. Missing neighbors, surface atoms are bound by weaker forces, so the vibration frequencies for these atoms are lower than those for interior atoms. Theoretical investigations have shown that surface phonons of microcrystals should be detectable by means of Raman scattering if the particles are small in comparison with the wavelength of the incident laser light.⁹ In this article we shall investigate the surface phonons of ZnS nanoparticles, using Raman scattering and infrared absorption.

II. EXPERIMENT

By using an evaporation technique combined with an induction heating method,¹⁰ fine Zn particles were prepared, and as-prepared Zn particles were mixed with 50 ml Na₂S solution at conditions of ultrasonic oscillation for 10 h. It is expected that a surface reaction will occur between Zn/ZnO (Zn coated with ZnO) particles and the Na₂S solution owing to the high reactivity of the fine Zn particles, and then ZnS nanoparticles will be formed. After the chemical reaction was completed, the crystals were washed by water for a long time so that Na ions in our sample were removed. Using an x-ray diffraction meter [(XRD), Rigaku D/Max-γC], the crystal structure of the sample was determined. The particle size was estimated by transmission electron microscopic [(TEM), JEM-200CX] observations. Raman scattering was performed in air ($\epsilon_m = 1$) using the 514.5 nm line of an Ar-ion laser (SPEX) as excitation light. Infrared absorption was measured on an infrared spectrophotometer (Nicolet 170

SX) in a medium of paraffin liquid ($\epsilon_m = 2.4$). All our measurements were carried out at room temperature.

III. RESULTS AND DISCUSSION

Figures 1(a) and 1(b) show the XRD before and after the sample was treated chemically in an Na₂S solution. Comparing Fig. 1(b) with Fig. 1(a), we find that a new peak centered at $2\theta = 28.68^\circ$ appears after being treated in the Na₂S solution, which corresponds to (111) of β -ZnS. Other diffraction peaks of ZnS are close to those of ZnO, they may be concealed in the diffraction peaks of ZnO. Therefore we conclude that the ZnS prepared is in zinc-blende structure. Figure 2 shows a typical TEM photograph of our sample. The black particles are Zn/ZnO and the white ones are ZnS. The size of ZnS particles can be estimated to be smaller than 10 nm.

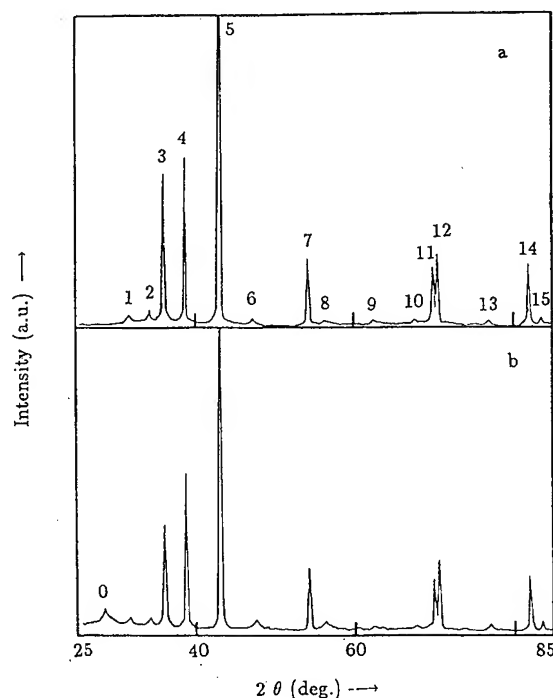


FIG. 1. XRD spectra: (a) before and (b) after the sample was treated in Na₂S solution (peaks 3–5, 7, and 11–15 for Zn; peaks 1,2,6,8–10 for ZnO; peak 0 for ZnS).

^{a)}Electronic mail: Postmaster@hndx.whnet.edu.cn

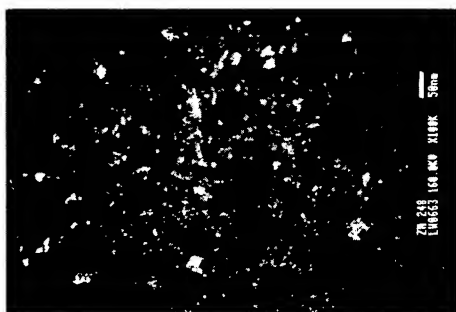


FIG. 2. A typical TEM photograph of the sample.

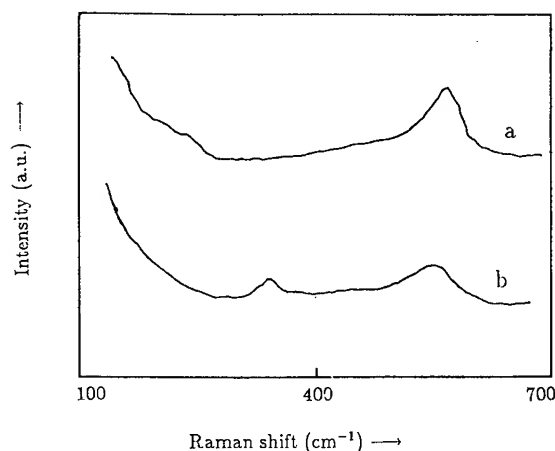
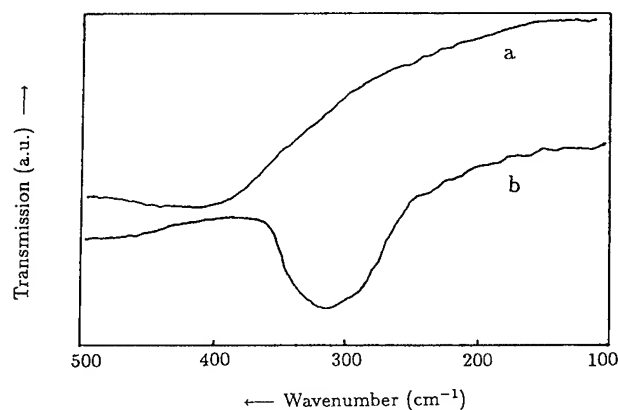
Figure 3 presents the Raman scattering spectra. We can see from Fig. 3(a) that there is only one broad peak around 561 cm^{-1} , which can be assigned to ZnO.¹¹ In Fig. 3(b) an additional peak centered at 332 cm^{-1} appears beside the original one.

Figures 4(a) and 4(b) show the far-infrared transmission spectra. Before the chemical reaction, the sample has no evident absorption in the range of $100\text{--}500\text{ nm}$. However, an intense absorption peak at 315 cm^{-1} can be found after the chemical treatment. It is believed that the Raman scattering peak at 332 cm^{-1} and the infrared absorption peak at 315 cm^{-1} come from β -ZnS prepared during the chemical reaction. However, they are not the fundamental vibrational modes of β -ZnS (see Table I). We note that they lie between the ω_{TO} and ω_{LO} , the long wavelength transverse and longitudinal optical phonon frequencies of ZnS. We will show that they are just the surface phonon mode of ZnS particles.

Consider a semiconductor sphere of radius R surrounded by a medium having a real, frequency-independent dielectric constant ϵ_m . Its undamped dielectric constant has the form¹²

$$\epsilon(\omega) = \epsilon_\infty \frac{\omega^2 - \omega_{\text{LO}}^2}{\omega^2 - \omega_{\text{TO}}^2}, \quad (1)$$

where ϵ_∞ is the dielectric constant at high frequency, and ω_{LO} and ω_{TO} are the LO and TO eigenfrequencies related by

FIG. 3. Raman spectra of the sample: (a) before and (b) after the sample was treated in Na_2S solution.FIG. 4. Far-infrared transmission spectra of the sample: (a) before and (b) after the chemical reaction in Na_2S solution.

$$\frac{\omega_{\text{LO}}^2}{\omega_{\text{TO}}^2} = \frac{\epsilon_0}{\epsilon_\infty}, \quad (2)$$

where ϵ_0 is the static dielectric constant. However, for the surface modes, the following equation must be satisfied:

$$\epsilon(\omega) = -\frac{l+1}{l} \epsilon_m, \quad l=1,2,3,\dots \quad (3)$$

Thus, combining Eqs. (1) and (3), the surface mode frequencies can be calculated. For small crystallites, the macroscopic theory has predicted that the major part of the scattering or absorption is from the $l=1$ mode.¹³ So only the first-order mode is concerned below. When $\epsilon_m=1$, corresponding to the Raman scattering case, the calculated surface mode frequency $\omega_s=332\text{ cm}^{-1}$, which is in agreement with the experimental result; when $\epsilon_m=2.4$, the calculated surface mode frequency $\omega_s=316\text{ cm}^{-1}$, which agrees with the infrared absorption result. Detail results are given in Table I. Thus, we conclude that the Raman and IR peaks observed experimentally are caused by surface phonon modes in ZnS nanoparticles.

It should be noted that, in our spectra, we have not found ZnS bulk vibration modes. In general, for sufficiently small particles, the contribution of the surface layers to the Raman scattering or infrared absorption is important because the surface to volume ratio is very large, and the number of atoms on the surface is comparable to or larger than that inside the particle. Thus it is not surprising that we cannot find bulk ZnS vibration modes in our sample.

TABLE I. Parameters ω_{TO} , ω_{LO} ^a and ϵ_0 ^b of ZnS used to calculate the surface mode frequencies ω_s , the calculated ω_s , and the experimental results.

ϵ_0	ω_{TO}	ω_{LO}	$\omega_s(\text{cm}^{-1})$			
			$\epsilon_m=1$		$\epsilon_m=2.4$	
			Calc.	Exp.	Calc.	Exp.
8.3	275	351	331	332	316	315

^aSee Ref. 14.

^bSee Ref. 15.

IV. SUMMARY

In summary, the following conclusions can be drawn:

- (1) ZnS nanoparticles can be prepared by slow surface reaction between fine Zn particles and Na_2S solution.
- (2) In both Raman and infrared spectra, a broad peak can be observed, which can be assigned to the surface phonon mode of ZnS nanoparticles.

ACKNOWLEDGMENTS

This work was supported by the State Major Basic Research Project "Nano-Materials Science" (85-6 NMS), the National Natural Science Foundation of China (NSFC), and the National Advanced Materials Committee of China (NAMC).

- ¹S. Schmitt-Rink, D. A. B. Miller, and D. S. Chemla, *Phys. Rev. B* **35**, 8113 (1987).
- ²M. Kumagai and T. Takagahara, *Phys. Rev. B* **40**, 12359 (1989).
- ³A. J. Shields, J. L. Osborn, M. Y. Simmons, M. Pepper, and D. A. Ritchie, *Phys. Rev. B* **52**, R5523 (1995).
- ⁴A. D. Yoffe, *Adv. Phys.* **42**, 173 (1993).
- ⁵V. Novotony and P. P. Meinke, *Phys. Rev. B* **8**, 4186 (1973).
- ⁶Ph. Buffat and J.-P. Borel, *Phys. Rev. B* **13**, 2287 (1976).
- ⁷M. Fujii, T. Nagareda, S. Hayashi, and K. Yamamoto, *Phys. Rev. B* **44**, 6243 (1991).
- ⁸Y. Sasaki and C. Horie, *Phys. Rev. B* **47**, 3811 (1993).
- ⁹R. Ruppini, *J. Phys. C* **8**, 1969 (1975).
- ¹⁰Jianfeng Xu, Tiejun Zhou, and Youwei Du, *Phys. Status Solidi A* **153**, 313 (1996).
- ¹¹Jianfeng Xu, Guangxu Cheng, and Youwei Du, *Chin. Phys. Lett.* **13**, 765 (1996).
- ¹²M. C. Klein, F. Hache, D. Richard, and C. Flytzanis, *Phys. Rev. B* **42**, 11123 (1990).
- ¹³R. Ruppini and R. Englman, *Rep. Prog. Phys.* **33**, 149 (1970).
- ¹⁴H. Kanie, M. Nagano, and M. Aoki, *Jpn. J. Appl. Phys.* **30**, 1360 (1991).
- ¹⁵N. F. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals*, 2nd ed. (Dover, New York, 1964).

Photovoltaic study of nanocrystalline TiO₂ film modified with dye molecules

Jinghuai Fang,^{a)} Jingwen Wu, Xiaomei Lu, Jianhua Gu, and Zuhong Lu

National Laboratory of Molecular and Biomolecular Electronics, Southeast University, Nanjing 210096, People's Republic of China

(Received 6 March 1997; accepted 10 March 1997)

The nanocrystalline TiO₂ film is modified simultaneously with porphyrin dye molecules and phthalocyanine dye molecules by chemical adsorption. Since porphyrin and phthalocyanine dye have different absorption character in the visible range, the TiO₂ electrode sensitized by a combination of porphyrin and phthalocyanine dye molecules not only obtains a better spectral match to the solar spectrum, but also the photoresponse of the electrode is extended to cover most of the visible range. This kind of electrode exhibits a significant improvement of energy conversion efficiency. © 1997 American Vacuum Society. [S0734-211X(97)01804-0]

I. INTRODUCTION

Photosensitization of wide-band-gap semiconductors such as TiO₂ by adsorbed organic dyes has become an important research field due to its potential applications in solar energy.¹⁻⁴ In the configuration, organic dye coated on the surface of the semiconductor is used to harvest incident light. Light absorption of the dye is followed by an electron injection from the excited state of the dye into the conduction band of the semiconductor. The remaining positive hole is scavenged by a redox couple in solution. The major advantages of the organic cells are the ease of fabrication and low cost. It is especially convenient to tailor properties of organic dye by the method of molecular design. Extensive research on the sensitization of solar cells has shown that the dye should be adsorbed on the semiconductor surface in a closely packed monolayer for maximum sensitization efficiency, but the light-harvesting efficiency of a single dye monolayer is very low. Thus, that the dye cells based on a plain electrode could not effectively utilize incident light is a major reason for its poor power conversion efficiency. Recently, Gratzel and co-workers⁵ developed a new kind of solar cell based on a highly porous TiO₂ substrate electrode with organic dye as a sensitizer. The internal surface area of these electrodes was so large that only a monolayer of dye is sufficient to absorb most of the incident light in its absorption spectral range. The energy conversion efficiency was improved.

Generally, the absorption of a single organic dye only covers part of the visible spectrum scope. For an efficient solar energy conversion, however, materials are needed to exhibit a high absorption coefficient over the whole spectral region from near-infrared to ultraviolet. Porphyrin and phthalocyanine dyes are very interesting photosensitization compounds. Their respective absorption is located in the different region of visible range. Meanwhile, they show a strong sensitization effect for the wide-band-gap semiconductor and extend the photoresponse of the wide-band-gap semiconductor electrode into the visible region. In this article, the microporous titanium dioxide (TiO₂) electrode has

been prepared on a conducting glass support. The electrode has such a high surface roughness factor that two different dye molecules could be coated simultaneously on the electrode surface, and that less than even a monomolecular layer of dye coating could absorb most of the incident light in the respective spectral range of dye absorbance. Compared with conventional electrodes using either porphyrin or phthalocyanine dye, a better spectral match to the solar spectrum was obtained for a microporous TiO₂ electrode sensitized by a combination of porphyrin and phthalocyanine dye molecules.

II. EXPERIMENT

A. Preparation of electrode

The nanocrystalline films were fabricated as follows. A colloidal solution was prepared by adding commercial TiO₂ power (P25, Degussa, a mixture of ~25% rutile and 75% anatase, surface area 55 m²/g, diameter 25 nm) to a small amount of water and surfactant (Triton X-100). The colloidal solution was then coated onto a conducting glass support. After air drying, the sample was heated in air in a Lindberg tube furnace in which the temperature was increased gradually to 450 °C and then kept at 450 °C for 30 min. After heat treatment, the TiO₂ film was strongly attached to the glass support.

The dyes of zinc porphine (ZnPP) and gallium phthalocyanine (GaPc) were coated onto the TiO₂ electrode by immersing TiO₂ electrode in a methanol solution with a respective dye concentration of about 10⁻³ mol/L. After at least 3 h of soaking, we obtained a dye modified TiO₂ electrode.

B. Experimental setup

The morphology of the TiO₂ electrode was examined with an atomic force microscope (DICO, NanoScope 3). The absorption spectra were recorded with a Shimadzu UV-2201 spectrometer. A two-electrode photoelectrochemical (PEC) cell, composed of a dye/TiO₂ electrode, a counterelectrode consisting of conducting glass sputter coated with a 1 μm Pt

^{a)}Electronic mail: jhfang@seu.edu.cn

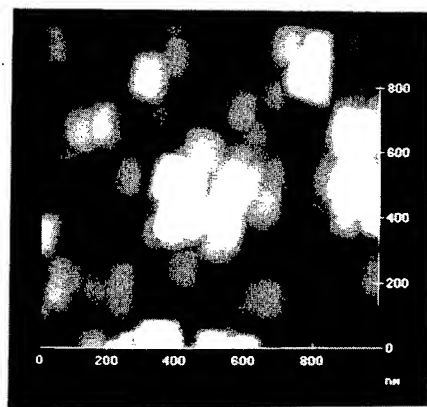


FIG. 1. The morphology of the TiO_2 microporous electrode. A typical operating force between the tip and TiO_2 sample surface is about 10^{-8} N.

film, and an electrolyte containing 0.1 M Na_2S , and 0.01 M Na_2SO_4 , was used for photovoltaic studies. The cell had an area of 0.5 cm^2 . The short-circuit photocurrent was measured with a Model CMBP-1 potentiostat. Monochromatic illumination was obtained by using a 500 W xenon arc lamp in combination with a grating monochromator, Model WPG3D. The light intensity was calibrated by using a model OM-1001C radiometer/photometer.

III. RESULTS AND DISCUSSION

A. AFM studies

The atomic force microscopy (AFM) image in Fig. 1 shows the morphology of the TiO_2 film on a conducting glass support. The TiO_2 film is microporous since it is composed of interconnected particles and pores. These pores not only increase the surface area and allow the electrode to adsorb more dye molecules, but also enable adsorbed dye molecules to contact directly with electrolyte. When dye molecules absorb incident light, the photogenerated electrons can transfer directly into TiO_2 , at the same time, the holes remaining in the dye can be rapidly extracted by electrolyte, decreasing the probability of recombination and increasing photoresponse. It can also be seen that the TiO_2 particles are uniform with an average diameter of about 70 nm, which is much larger than the size of the TiO_2 particles in the colloidal solution. Experiment results indicate that the heating process is crucial to the morphology of the TiO_2 electrode. The size of the particles in the TiO_2 film increases with the rising of temperature.

B. Absorption characteristics and photovoltaic study

The absorption spectra of the TiO_2 electrode with and without a coating of ZnPP or GaPc have been measured and are shown in Fig. 2. A conducting glass support is used as a reference in the measurement. Since the TiO_2 adopted in our experiment contains 25% rutile and 75% anatase, and the band gap of rutile is 3 eV, whereas that of anatase is 3.2 eV, corresponding to a fundamental absorption edge of 413 and 388 nm, respectively. So the bare TiO_2 film exhibits the fun-

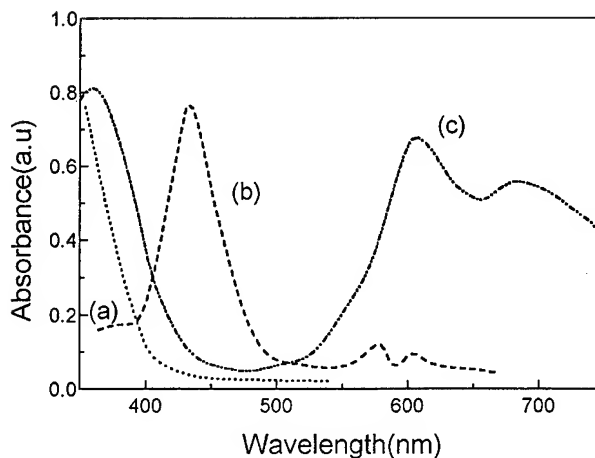


FIG. 2. Absorption spectra of bare TiO_2 electrode (a), ZnPP/ TiO_2 electrode (b), and GaPc/ TiO_2 electrode (c).

damental absorption edge of rutile. Curve (b) is the absorption spectrum of ZnPP/ TiO_2 electrode. It shows that the ZnPP/ TiO_2 electrode displaces the characteristic absorption of ZnPP with a maximum in the visible region at around 430 nm. Curve (c) is the absorption spectrum of the GaPc/ TiO_2 electrode. It indicates that sensitization of the TiO_2 electrode with GaPc dye molecules could extend absorbance of the electrode into the red region. It can also be seen that the absorption of the ZnPP/ TiO_2 electrode and GaPc/ TiO_2 electrode are distributed in different regions of the visible range. So the absorption spectrum of the comodified TiO_2 electrode with ZnPP and GaPc dye molecules can cover most of the visible light. In addition, further experimental results show that the light absorption of the dye sensitized TiO_2 electrode increases with the increase of the thickness of the TiO_2 film, we can infer from the above that dye molecules can penetrate into pores of the TiO_2 electrode to be modified on the internal particles surface of the TiO_2 electrode and thus the number of the dye molecules directly on the surface of TiO_2 electrode increases.

Figure 3 shows the absorption spectrum [curve (b)] and

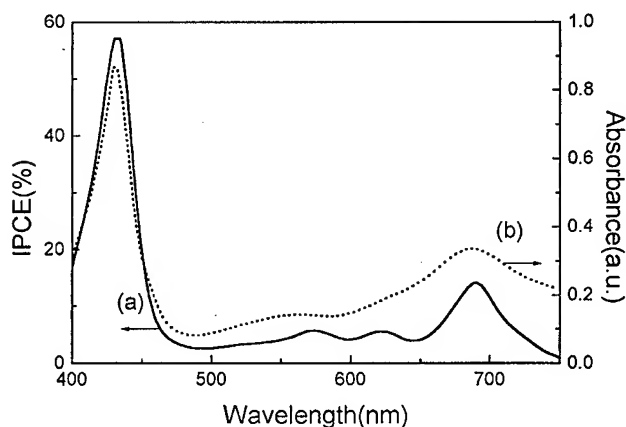


FIG. 3. Absorption spectrum (b) and photocurrent action spectrum (a) of ZnPP-GaPc/ TiO_2 .

photocurrent action spectrum [curve (a)] of the ZnPP–GaPc/TiO₂ electrode. It is obvious from curve (b) that the absorption of the TiO₂ electrode modified with ZnPP and GaPc dye molecules is wider than that of TiO₂ electrode modified with single dye molecules, and can match the solar spectrum. The photocurrent action spectrum of the ZnPP–GaPc/TiO₂ electrode correlates well with its absorption spectrum, indicating the ZnPP and GaPc have a significant contribution to the photocurrent of the electrode. The incident-photon-to-current-conversion (IPCE) is 57% at 430 nm. Considering the light reflection loss, an IPCE up to 70% is reached. The high yield in the IPCE of the PEC cell based on the TiO₂ microporous electrode can be explained as follows: The dye is absorbed on the surface of the TiO₂ electrode, due to the large surface area factor of the microporous electrode, only the first monolayer of adsorbed dye is sufficient to absorb the incident light nearly completely, the photogenerated carriers resulted from absorption of incident photons by dye molecules can be easily transported into the conduction band of the TiO₂ semiconductor, improving the charge separation efficiency.

IV. CONCLUSION

As a summary, the microporous TiO₂ electrode is of large surface roughness factor, we can use simultaneously two different dye molecules to modify TiO₂ electrode. Compared with conventional electrodes using either porphyrin or phthalocyanine, the cosensitization of TiO₂ electrodes with ZnPP and GaPc efficiently extends the optical absorption spectrum and photocurrent spectrum into a wider visible range. Therefore the cosensitization with different dye molecules is an effective way to improve the photovoltaic properties of dye sensitized cells.

¹T. Miyasaka, T. Watanabe, A. Fujishima, and K. Honda, *Nature (London)* **638**, 277 (1979).

²M. Krishnan, X. Zhang, and A. J. Bard, *J. Am. Chem. Soc.* **371**, 106 (1984).

³F. Willing, R. E. Eichberger, N. S. Sundaresan, and B. A. Parkison, *J. Am. Chem. Soc.* **2702**, 112 (1990).

⁴J. Segui, S. Hotchandani, D. Baddou, and R. M. Leblance, *J. Phys. Chem.* **8807**, 95 (1991).

⁵B. O. Regan and M. Gratzel, *Nature (London)* **737**, 253 (1991).

Visible luminescence from Si/SiO₂ superlattices

S. V. Novikov,^{a)} J. Sinkkonen, and O. Kilpelä

Electron Physics Laboratory, Helsinki University of Technology, Otakaari 7A, FIN-02150 Espoo, Finland

S. V. Gastev

A.F. Ioffe Institute, 194021 St.-Petersburg, Russia

(Received 6 March 1997; accepted 10 March 1997)

Si/SiO₂ superlattices were grown by molecular beam epitaxy using *in situ* oxidation by the rf-plasma source. Transmission electron microscopy shows high uniformity of the grown layers. Optical band gap obtained from the absorption measurement changes from 1.2 eV for thick samples to 2.3 eV for thin samples, indicating quantum confinement in the Si layers. Room temperature photoluminescence (PL) was observed in the spectral range 1.9–2.3 eV. Two different mechanisms of PL were found. In the Si layer thickness range 1–2 nm the luminescence arises from the quantum confined energy states in the ultrathin silicon layers. For Si layers with thickness more than 3 nm oxygen related defects are the main origin of PL. © 1997 American Vacuum Society. [S0734-211X(97)01404-2]

I. INTRODUCTION

Luminescent materials are widely used by the industry in optoelectronic devices. Until recently the semiconductor industry has concentrated most of its resources in the field of the direct band gap compounds, such as III–V compounds, as the direct band gap results in strong optical absorption and efficient light emission. Today though, in a search for materials which are more compatible with existing silicon technologies than the III–V compounds, the industry has started to research the possibility of exploiting the luminescent properties of silicon and silicon-based compounds.¹ These silicon-based materials would be fully compatible with the existing silicon process technology and would allow the integration of optoelectronics with other silicon-based devices.

Crystalline silicon itself shows very poor luminescence but experimental results have led to the belief that a number of various types of silicon and silicon-based compounds might have useful luminescent properties. As an example, discovery of the bright luminescence from electrochemically etched porous silicon² spurred increased research efforts aimed at fabricating efficient silicon-based light-emitting diodes (LEDs). However, because the emission wavelength of this material does not show size dependence typical for quantum confinement, the most recent model of this luminescence supposes the existence of oxygen-related surface localized defects which take part in the light emission.^{3,4} Recently the luminescence from Si/SiO₂ superlattices, grown by molecular beam epitaxy (MBE) and *ex situ* oxidation of thin Si layers was reported.⁵ Experimental results have shown that the dependence of the light emission wavelength and band gap with Si layer thickness can be explained in terms of quantum confinement of the electrons in the two-dimensional (2D) silicon layers. However, *ex situ* oxidation does not take full advantage of MBE's capability for ultraclean processing, due to the possible contamination of the surface during exposure of the sample to the atmosphere. Here we report the fabrication of fully *in situ* (MBE) grown Si/SiO₂ superlat-

tices, which have luminescence in the visible range. Using *in situ* rf-plasma source for oxidation allows one to eliminate the waiting time for loading the wafers into the deposition chamber after *ex situ* oxidation. It also offers more precisely controlled conditions for the oxide growth.

II. EXPERIMENT

A modified VG Semicon V80M MBE system is used for growth of the Si/SiO₂ superlattices. Oxygen (O₂) is used as the source gas. Gas fluxes are controlled by mass flow controllers. A plasma activated gas cell (Oxford Applied Research MPD-21) is used as the oxygen source. The Leybold-Heraeus 180° electron gun with the quartz deposition rate controller is used for Si deposition. The substrate temperature is monitored by an optical pyrometer and a thermocouple. The pumping system of the growth chamber consists of an oil diffusion pump, ion pump, and titanium sublimation pump. The growth chamber is also equipped with a reflection high-energy electron diffraction (RHEED) system for monitoring the growth.

Quarters of 3 in. silicon (100) wafers are used as substrates. The Shiraki⁶ chemical treatment is used as a chemical cleaning process. The thin native oxide is evaporated at 780 °C under the Si flux of 0.01 nm/s. Once the 2×1 Si surface superstructure appears, the Si beam is interrupted. Then a 50-nm-thick Si layer is grown with a rate 0.07 nm/s at 500 °C. Finally the substrate is annealed 15 min at 820 °C. This cleaning procedure allows one to obtain a very smooth surface.⁷ The surface roughness measured by atomic force microscopy (AFM) is less than 2 nm at 1 μm scan. Then the Si surface is oxidized at room temperature. Oxygen mass flow rate is 4.0 sccm during the oxidation which roughly corresponds to the pressure 3×10⁻⁵ mbar inside the growth chamber. Typical power input for the rf coil is 260 W. The growth time is 10 min which corresponds to the SiO₂ thickness of 1.0 nm. The amorphous silicon layers with thicknesses of 1.5–4.0 nm are deposited at room temperature with the growth rate 0.07 nm/s on such an SiO₂ layer. The layers for transmission measurements are grown at the same condi-

^{a)}Corresponding author.

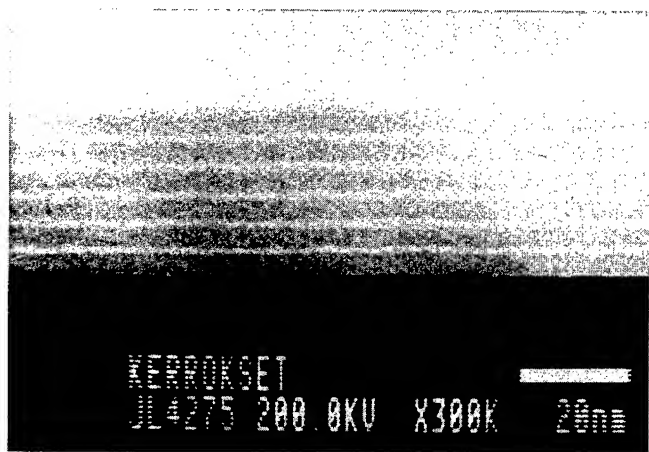


FIG. 1. Cross-sectional bright field transmission electron micrograph of the superlattice which consists of six Si layers (dark bands) and seven SiO₂ layers (light bands).

tions on the polished fused silica substrate. The samples have been characterized using ellipsometry, RHEED, transmission electron microscopy (TEM), AFM, optical absorption spectroscopy, and photoluminescence (PL).

III. RESULTS AND DISCUSSION

The TEM cross-sectional image of a six-period Si/SiO₂ superlattice by a Jeol 2010 200 kV electron microscope is shown in Fig. 1. The Si (4.0 nm) and SiO₂ (1.0 nm) layers correspond to white and dark regions, respectively. The layers are smooth and uniform in their thickness. The interface between Si and SiO₂ is abrupt.

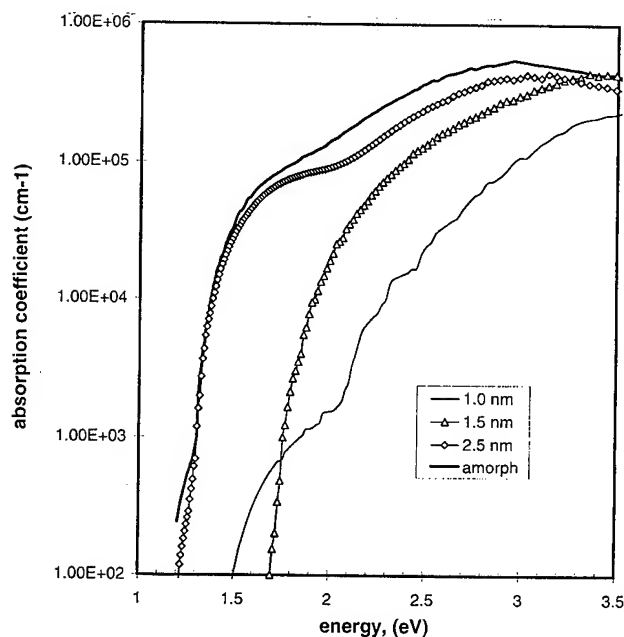


FIG. 2. Absorption spectra of the Si/SiO₂ superlattices grown on the fused silica substrate with different thickness of *a*-Si layers (the thickness is shown in the inset). The thickness of SiO₂ layers is 1 nm.

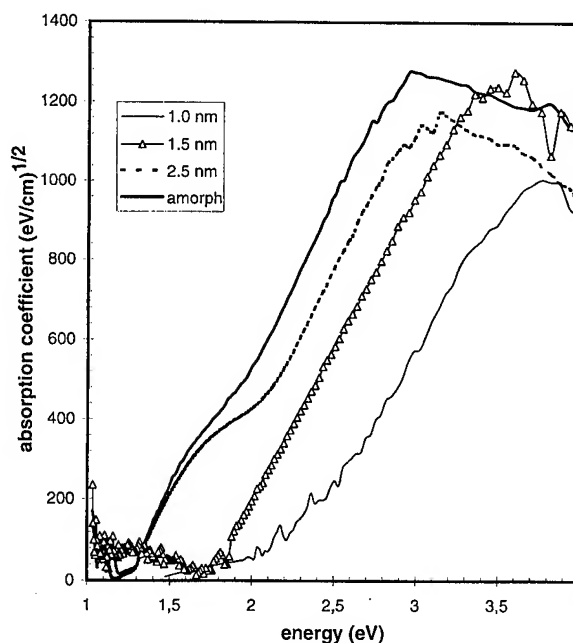


FIG. 3. $(\alpha h\nu)^{1/2} - h\nu$ absorption spectra of the Si/SiO₂ superlattices grown on the fused silica substrate with different thickness of *a*-Si layers (the thickness is shown in the inset). The thickness of SiO₂ layers is 1 nm.

Absorption spectra of the superlattices with different silicon layer thicknesses, grown on fused silica substrate, are shown in Fig. 2. It is seen that the absorption edge position is moving to short wavelength when silicon layer thickness decreases. The optical band gap is determined from the deduced $\alpha(h\nu)$ data using the extrapolation of $(\alpha h\nu)^{1/2}$ curve to zero ordinate, i.e., the constant-*n* form of the Tauc law:⁸

$$(\alpha h\nu)^{1/2} = \text{const} \cdot (h\nu - E_g). \quad (1)$$

The absorption spectra in $(\alpha h\nu)^{1/2} - h\nu$ coordinates are shown in Fig. 3. It is seen that the optical band gap determined from this spectra changes from 1.2 eV for thick samples to 2.3 eV for the sample with 1 nm thickness. According to Ref. 4, the energy gap *E* for 1D confined Si is given by

$$E(\text{eV}) = 1.60 + 0.72d^{-2}, \quad (2)$$

where *d* is the amorphous silicon layer thickness in nm. However, in our case the absorption edge position is in good agreement with the formula

$$E(\text{eV}) = 1.2 + 0.72d^{-2}. \quad (3)$$

The value of the band gap energy (1.2 eV) obtained in this case is less than usually reported for the amorphous silicon.

The room temperature PL spectra of superlattices are measured using 488 nm (2.5 eV) argon laser excitation, double monochromator, and photomultiplier. The PL spectra from the Si/SiO₂ superlattices grown on silicon and fused silica substrate are shown in Figs. 4 and 5, respectively. For the Si layer thicknesses more than 2.5 nm our structures have a broad PL peak at 650 nm which corresponds to the energy 1.9 eV. According to Eq. (2) the influence of the quantum

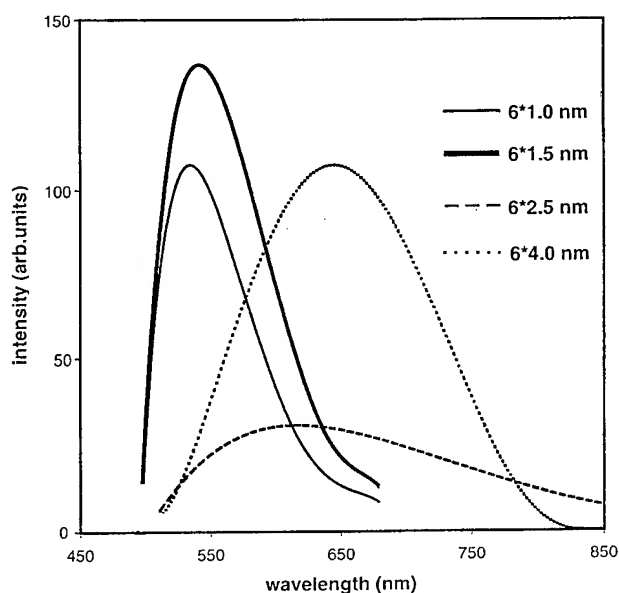


FIG. 4. PL spectra of the Si/SiO₂ superlattices grown on silicon substrate with different thickness of *a*-Si layers (the thickness is shown in the inset). The thickness of SiO₂ layers is 1 nm.

confinement on the peak energy in this thickness range is very weak. The PL intensity depends on the total Si thickness in the superlattice. Thus in this case the luminescence is a property of the silicon layer. However, the value of the PL peak energy position (1.9 eV) is much higher than that estimated by using absorption edge measurement data in Eq. (3).

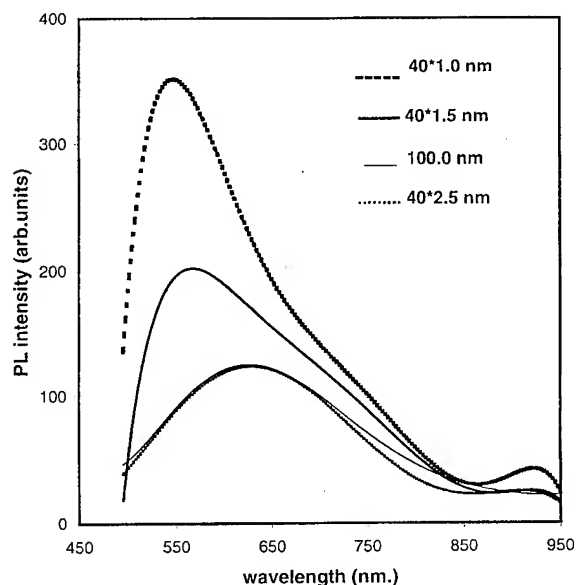


FIG. 5. PL spectra of the Si/SiO₂ superlattices grown on fused silica substrate with different thickness of *a*-Si layers (the thickness is shown in the inset). The thickness of SiO₂ layers is 1 nm.

The luminescence peak position is very similar to ones for porous silicon.⁹ It was reported in Refs. 3 and 4 that oxygen related defects such as nonbridging oxygen hole centers (NBOHCS) are the main source of PL in porous silicon. Our samples may have a higher oxygen content in the silicon layers than reported in Ref. 4, because the silicon growth and the oxidation were carried out in the same chamber with a 30–60 min waiting between oxidation and Si growth. In Ref. 5 no PL was reported from superlattices with a Si layer thickness more than 3 nm. Therefore we suppose that the oxygen related defects are the main source of PL from Si/SiO₂ superlattices in this thickness range.

In the silicon layer thickness range 1–2 nm the luminescence peaks are blueshifting up to 530 nm (2.3 eV) as shown in Figs. 4 and 5. The peak energy shift is consistent with Eq. (2). The PL intensity is higher than in thickness range 2–4 nm, although the total Si thickness decreases. This might stem from the enhancement of the oscillator strength with the decrease of the Si layer thickness owing to the quantum confinement.¹⁰ Thus we are observing quantum confinement of the energy states. The photoluminescence spectra have a long wavelength tail at the 600–700 nm range, which corresponds the PL from oxygen related centers, but its intensity is low because the total silicon thickness in the superlattice in this case is small.

IV. CONCLUSION

In conclusion, the room temperature PL in the range 540–650 nm was observed from the Si/SiO₂ superlattices grown using MBE and *in situ* oxidation. The oxygen related defects in silicon are the main origin of PL from superlattices with a silicon thickness of more than 2.0 nm. In the Si layer thickness range of 1–2 nm the luminescence arises from the quantum confined energy states in the ultrathin silicon layers.

ACKNOWLEDGMENT

The authors would like to thank Dr. J. Laurila for the TEM analysis.

¹A. B. Miller, *Nature (London)* **378**, 238 (1995).

²L. T. Canham, *Appl. Phys. Lett.* **57**, 1046 (1990).

³L. N. Skuja and A. R. Silin, *Phys. Status Solidi A* **56**, K11 (1979).

⁴S. M. Prokes, *IEEE J. Sel. Topics Quantum Electron.* **1**, 1140 (1995).

⁵Z. H. Lu, D. J. Lockwood, and J.-M. Baribeau, *Phys. Rev. Lett.* **76**, 539 (1996).

⁶A. Ishizaka and Y. Shiraki, *J. Electrochem. Sci. Technol.* **150**, 966 (1986).

⁷J. Lundsgaard Hansen, S. Yu. Shiryayev, and E. V. Thomsen, *J. Cryst. Growth* **157**, 317 (1995).

⁸T. Tiedje, B. Abeles, P. D. Persans, B. G. Brooks, and G. D. Cody, *J. Non-Cryst. Solids* **66**, 345 (1984).

⁹D. W. Cooke, B. L. Bennet, E. H. Farnum, W. L. Hults, K. E. Sickafus, J. F. Smith, J. L. Smith, T. N. Taylor, and P. Tiwari, *Appl. Phys. Lett.* **68**, 1663 (1996).

¹⁰J. B. Khurgin, E. W. Forsythe, G. S. Tompa, and B. A. Khan, *Appl. Phys. Lett.* **69**, 1241 (1996).

Numerical study of optical transmission in sandwich slabs: Implication to photon scanning tunneling microscopy

Shu Wang, Mufei Xiao, and Jesús M. Siqueiros

Instituto de Física, Universidad Nacional Autónoma de México, Apartado Postal 2681, Ensenada, Baja California 22800, México

(Received 8 September 1996; accepted 1 April 1997)

In this work, we present a numerical simulation of scanning near-field optical microscopy. The simulation is based on a recently developed macroscopic theory which rigorously solves the electromagnetic field in a two-dimensional dielectric-air-dielectric sandwich system with arbitrary one-dimensional structures at its two interfaces. A tiny dielectric triangular is added on one interface to represent the probe tip in the device, while on the other interface, two identical topographic defects with various profiles are attached to represent the sample surface. The probe tip is assumed scanning at a constant height above the sample surface which is illuminated by a p -polarized light. The intensity of the local field averaged along a cross section inside the tip is calculated to represent the signal level of the microscope. The feasibility of the theoretical scheme is numerically demonstrated, and various aspects in the microscopy, such as the effects of the separation of the two objects, and of the shape differences between ridges and grooves are discussed. © 1997 American Vacuum Society. [S0734-211X(97)02704-2]

I. INTRODUCTION

Scanning near-field optical microscopy (SNOM)^{1,2} has attracted much attention in the past ten years. Numerous theories for SNOM have been reported, both in microscopic and macroscopic approaches.³⁻⁶ A useful type of the near-field optical microscope which allows imaging of transparent samples is named the photon scanning tunneling microscope (PSTM). In this technique, the incident field is propagating inside the sample while the evanescent field generated by total internal reflection is collected by an uncoated optical fiber in the air side.²

There are several theoretical works^{7,8} devoted to PSTM based on the Rayleigh hypothesis, which assumes that the electromagnetic fields can be represented by the superposition of plane waves and thus the resulting expression can be used to satisfy the appropriate boundary conditions.⁹ However, these works are using a perturbative method, which can not be applied to the case where the resonance of the system becomes important. In a recent report,¹⁰ we proposed a new approach based on the application of the Fourier transform technique in PSTM. Our theory is also in the framework of the Rayleigh approach but we solve the inverse Fourier transform exactly, instead of perturbatively⁸ or iteratively.¹¹ The theory enables us to handle surface features in arbitrary shapes, sizes, and distance between scatterers.

This work is devoted to the numerical application of the developed theory¹⁰ in the PSTM configuration. We have calculated the averaged field intensity at the inside of the probe tip scanning at a constant height above a surface with two identical topographic defects. The theoretical scheme is numerically demonstrated for a p -polarized incident beam. Various aspects in the microscopy, such as the effects of the separation of the two objects, and of the shape differences between ridges and grooves are discussed.

The outline of this article is the following. In Sec. II, we

introduce the theory on which the calculations presented in this article are based. The numerical results obtained by this approach and the discussions are presented in Sec. III. Finally, we present the conclusions drawn from these results.

II. THEORY

The geometry used in this article is similar to that in Ref. 10, where a central medium of thickness D with dielectric constant ϵ_2 is sandwiched between a substrate (sample) of dielectric constant ϵ_1 and a superstrate (probe tip) of dielectric constant ϵ_3 . The dielectric constants being, in general, complex numbers. The z axis in our coordinate system is chosen to be perpendicular to the mean plane of the interfaces, therefore the x and y axis are parallel to these interfaces.

We assume a one-dimensional function to represent the tip profile in the z direction on the tip-air interface, i.e.,

$$z = h'g(x). \quad (1)$$

The air-sample interface is described by another function:

$$z = hf(x) - D. \quad (2)$$

h' and h in Eqs. (1) and (2) are constant. The functions $g(x)$ and $f(x)$ can be arbitrarily defined, provided the range of the validity of the Rayleigh hypothesis is satisfied.⁹

Let the sample side (medium 1) be illuminated by a laser beam. The incident beam can be described by a monochromatic plane wave of frequency ω . If the plane of incidence is chosen to be the x - z plane, the y component of the field is

$$U_i(x, z) = Ie^{i(Q_0x + \beta_0z)}, \quad z \leq hf(x) - D, \quad (3)$$

where I represents the amplitude of the electric field E_y in the case of s polarization, or the magnetic field H_y in the case of p polarization, Q_0 and β_0 are the x component and the z component of the incident wave vector, respectively.

The scattered fields (y component) can also be expanded into plane waves with all values of the wave vector component Q parallel to the surface:

$$U_t(x, z) = \int_{-\infty}^{\infty} dQ T(Q) e^{i(Qx + \beta_3 z)}, \quad z \geq h'g(x), \quad (4)$$

$$U_v(x, z) = \int_{-\infty}^{\infty} dQ [A(Q) e^{i\beta_2 z} + B(Q) e^{-i\beta_2 z}] e^{iQx}, \quad h'g(x) \geq z \geq hf(x) - D, \quad (5)$$

$$U_r(x, z) = \int_{-\infty}^{\infty} dQ R(Q) e^{i(Qx - \beta_1 z)}, \quad z \leq hf(x) - D, \quad (6)$$

where $U_t(x, z)$ is the transmitted field in medium 3, $U_v(x, z)$ is the field in medium 2, $U_r(x, z)$ is the reflected field in medium 1, and

$$\beta_j = \left(\epsilon_j \frac{\omega^2}{c^2} - Q^2 \right)^{1/2}, \quad j = 1, 2, 3,$$

are the perpendicular components of the wave vectors in the three media. In the Rayleigh hypothesis, the above expressions are continued in the boundary regions.

In the following, we write the boundary conditions:

$$U_v|_{z=h'g(x)} = U_t|_{z=h'g(x)}, \quad (7)$$

$$\frac{1}{\nu_2} \frac{\partial}{\partial n} U_v|_{z=h'g(x)} = \frac{1}{\nu_3} \frac{\partial}{\partial n} U_t|_{z=h'g(x)}, \quad (8)$$

$$(U_i + U_r)|_{z=hf(x)-D} = U_v|_{z=hf(x)-D}, \quad (9)$$

$$\frac{1}{\nu_1} \frac{\partial}{\partial n} (U_i + U_r)|_{z=hf(x)-D} = \frac{1}{\nu_2} \frac{\partial}{\partial n} U_v|_{z=hf(x)-D}, \quad (10)$$

where

$$\frac{\partial}{\partial n} \psi \equiv \hat{n} \cdot \nabla \psi, \quad (11)$$

\hat{n} is a unit vector perpendicular to the interfaces at an arbitrary point, $\nu_j = 1$ for s polarization and $\nu_j = \epsilon_j$ for p polarization.

The Eqs. (7)–(10) can be written as a matrix of the form

$$\int_{-\infty}^{\infty} dQ \mathcal{M}(Q, x) \mathcal{R}(Q) e^{iQx} = \mathcal{T}(Q_0, x) e^{iQ_0 x}, \quad (12)$$

where

$$\mathcal{M}(Q, x) = \begin{pmatrix} -e^{i\beta_3 g h'} & e^{i\beta_2 g h'} & e^{-i\beta_2 g h'} & 0 \\ \frac{Q \dot{g} h' - \beta_3}{\nu_3} e^{i\beta_3 g h'} & \frac{-Q \dot{g} h' + \beta_2}{\nu_2} e^{i\beta_2 g h'} & \frac{-Q \dot{g} h' - \beta_2}{\nu_2} e^{-i\beta_2 g h'} & 0 \\ 0 & -e^{i\beta_2 (f h - D)} & -e^{-i\beta_2 (f h - D)} & e^{-i\beta_1 (f h - D)} \\ 0 & \frac{Q \dot{f} h - \beta_2}{\nu_2} e^{i\beta_2 (f h - D)} & \frac{Q \dot{f} h + \beta_2}{\nu_2} e^{-i\beta_2 (f h - D)} & \frac{-Q \dot{f} h - \beta_1}{\nu_1} e^{-i\beta_1 (f h - D)} \end{pmatrix}, \quad (13)$$

$$\mathcal{R}(Q) = \begin{pmatrix} T(Q) \\ A(Q) \\ B(Q) \\ R(Q) \end{pmatrix}, \quad (14)$$

and

$$\mathcal{T}(Q_0, x) = \begin{pmatrix} 0 \\ 0 \\ -1 \\ \frac{Q_0 \dot{f} h - \beta_0}{\nu_1} \end{pmatrix} I e^{i\beta_0 (f h - D)}. \quad (15)$$

Our task is to determine the field amplitudes $T(Q)$, $A(Q)$, $B(Q)$, and $R(Q)$. First, we transform the integral in Eqs. (12)–(15) into a supermatrix where the wave number Q increases discretely, i.e.,

$$Q_n = Q_0 + 2\pi n/d, \quad n = \dots, -1, 0, 1, \dots, \quad (16)$$

where d is a distance segment where the $g(x)$ and $f(x)$ are defined.

With the above discrete wave number the various integral functions and matrices can be replaced by summations as¹⁰

$$\mathcal{R}(Q) = \sum_{-\infty}^{\infty} \mathcal{R}_n(Q_n) \delta(Q - Q_n). \quad (17)$$

Substituting Eq. (17) into Eq. (12), we get

$$\sum_{-\infty}^{\infty} \mathcal{M}(Q_n, x) \mathcal{R}_n(Q_n) e^{iQ_n x} = \mathcal{T}(Q_0, x) e^{iQ_0 x}. \quad (18)$$

If we multiply every one of the four rows of Eq. (18) by $e^{(-iQ_m x)}$ and integrate over d , we obtain

$$\sum_{-\infty}^{\infty} \hat{\mathcal{M}}(Q_n, m-n) \mathcal{R}_n(Q_n) = \hat{\mathcal{T}}(Q_0, m), \quad (19)$$

where every element of the matrices $\hat{\mathcal{M}}(Q_n, m-n)$ and $\hat{\mathcal{T}}(Q_0, m)$ is the Fourier coefficient of the corresponding matrix element of \mathcal{M} and \mathcal{T} . Apparently, when $d \sim \infty$ and $m \sim \infty$, one has the rigorous integration. In practice, in a

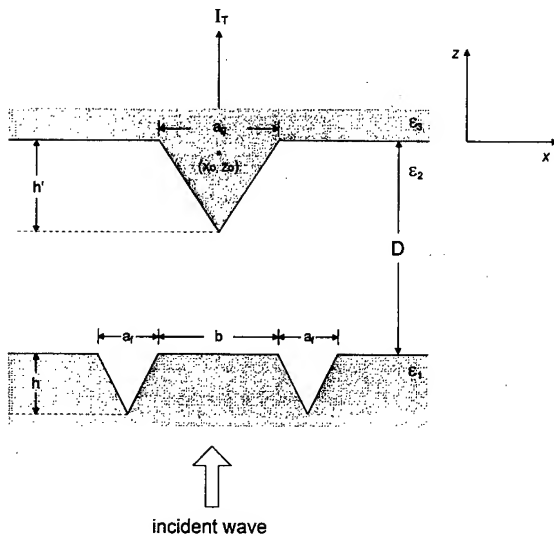


FIG. 1. A scheme showing the tip-sample conjunction. An air gap of separation $D=60$ nm and dielectric constant $\epsilon_2=1$ is sandwiched, on top, by a probe tip of ϵ_3 , and on the bottom, by substrate of ϵ_1 with subwavelength defects. The light comes from the substrate side. The groove width and depth are a_f and h , respectively. The tip width and height are a_g and h' , respectively. The two identical grooves are separated by a distance of b .

system in PSTM (like that in Fig. 1), the condition $d \sim \infty$ can be approximately replaced by $d \gg \text{wavelength} > a$, where a is the object spacing. Notice that T_n is the first element of $\mathcal{R}_n(Q_n)$, which is the amplitude of the transverse field trans-

mitted at order n . For a high-frequency sample ($a \ll c/\omega$) with small details, detection must take place in the near field and high harmonics are necessary. The maximum values of n and m can be defined, in computation practice, when the numerical results converge. Then the y component of the transmitted magnetic fields for p -polarization (if $\nu=\epsilon_1$) or electric fields for s polarization (if $\nu_j=1$) can be written as

$$U_t(x, z) = \sum_{-\infty}^{\infty} T_n e^{i[Q_n x - \beta_3(Q_n) z]}. \quad (20)$$

One can see that when $n \rightarrow \infty$, Eq. (20) leads to Eq. (4). The detected signal inside the tip is assumed proportional to the square modulus of the transmitted fields U_t .

III. NUMERICAL SIMULATION AND DISCUSSIONS

In this section we present some primary numerical simulation based on the formalism outlined in the preceding section. In these calculations, the distance segment is chosen as $d=10\lambda$. The numerical results converge at $n=m=30$. The calculations were performed on a Cray supercomputer.

As we have shown in Fig. 1, a probe tip (with triangular shape) is added to the upper medium with optical constant $n_3 = \sqrt{\epsilon_3} = 1.456$, which is equivalent to the optical constant of an optical fiber. On the opposite side there is a dielectric substrate with two subwavelength objects (triangular grooves). We assume that a p -polarized monochromatic plane wave of wavelength $\lambda=650$ nm is incident normally

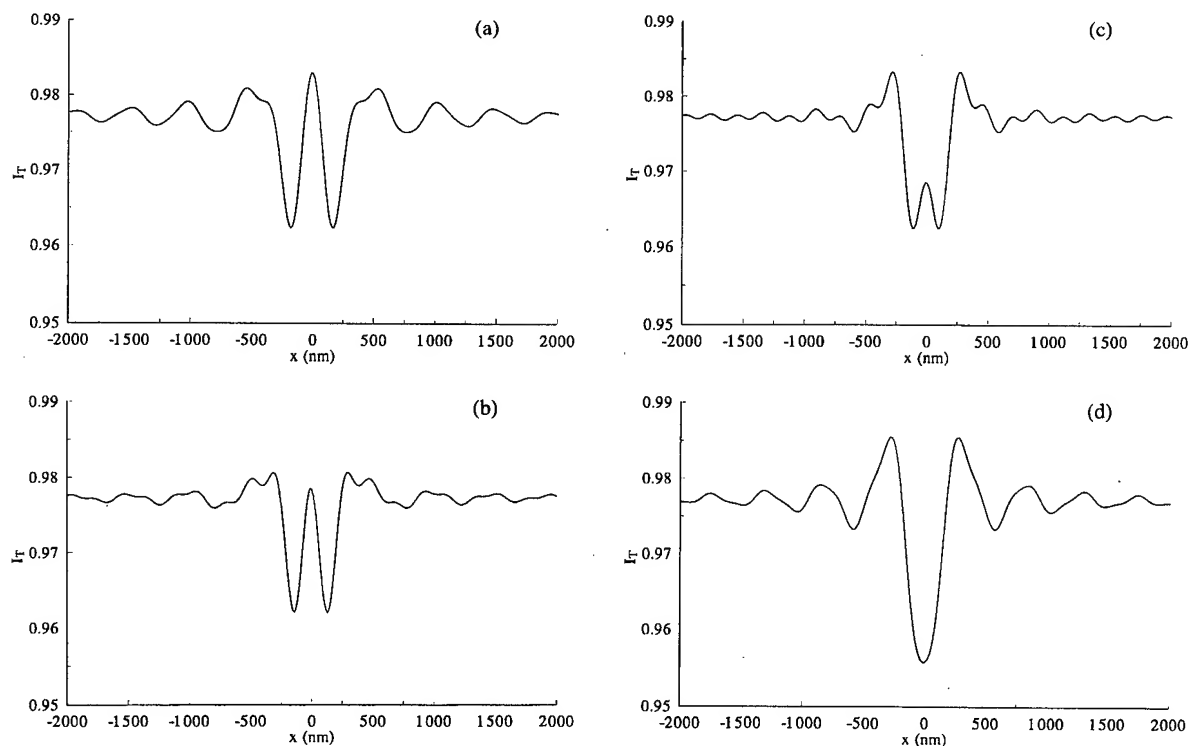


FIG. 2. Normalized near-field intensity I_T in transmission mode, scanning at constant height above the dielectric surface with two grooves. The probe width and height are $a_g=50$ nm and $h'=30$ nm, respectively. The groove width and depth are $a_f=80$ nm and $h=20$ nm, respectively. The separation of the grooves is: (a) $b=260$ nm, (b) $b=195$ nm, (c) $b=130$ nm, and (d) $b=65$ nm.

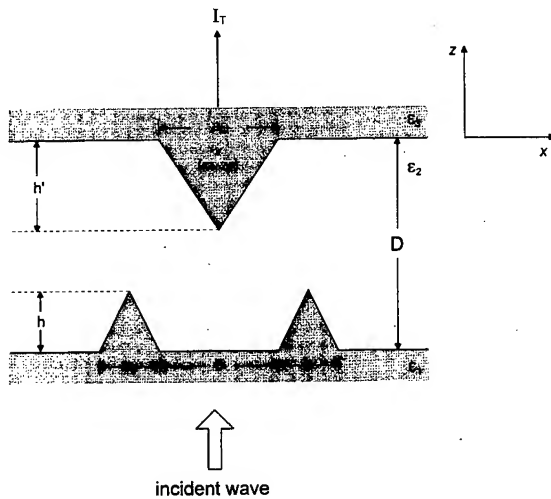


FIG. 3. Same as Fig. 1, but the grooves are replaced by ridges.

from the substrate side (medium 1). The plane of incidence is chosen to be in the x - z plane. We have calculated the field intensity $I_T = |\mathbf{B}_t|^2 / |\mathbf{B}_i|^2$ by averaging along a cross section ($z_0 = -10$ nm) inside the tip, where \mathbf{B}_t is the transmitted magnetic field and \mathbf{B}_i is the incident magnetic field.

We show in Figs. 2(a)–2(d) the field intensity I_T as the tip scans along a line at a constant height ($D = 60$ nm) above the dielectric surface (Fig. 1). There are two triangular shaped grooves separated by a distance of (a) $b = 260$ nm, (b) $b = 195$ nm, (c) $b = 130$ nm, and (d) $b = 65$ nm, respec-

tively. The probe width and height are $a_g = 50$ nm and $h' = 30$ nm, respectively. The groove width and depth are $a_f = 80$ nm and $h = 20$ nm, respectively. The optical constant of the medium where the incident light is coming from is $n_1 \equiv \sqrt{\epsilon_1} = 1.456$, that is, the optical constant of glass.

We see that the positions of the grooves (each of sub-wavelength size) are shown in the near-field image of the transmission mode by two dips at the same position. In the case of $b = 260$ nm [see Fig. 2(a)], the separation of these two dips shows very clearly on the image. When the two grooves approach each other, the features of the image get closer [see Figs. 2(b) and 2(c)]. This groove-dip correlation is lost when $b = 65$ nm, showing a strong dip at the symmetry center of the near-field image [see Fig. 2(d)]. This implies that the resolution of our simulated near-field optical microscope is about $b \sim 0.2\lambda$. If the tip-sample distance D increases, the calculated resolution will be lower, although we have not shown it explicitly in this work.

A different configuration is shown in Fig. 3. All parameters are the same as in Fig. 1, but we have replaced grooves by ridges.

Figures 4(a)–4(d) show the near-field intensity I_T as the tip scans along a line at a constant height $D = 60$ nm, above the dielectric surface (Fig. 3). The two triangular shaped ridges are separated by a distance of (a) $b = 260$ nm, (b) $b = 195$ nm, (c) $b = 130$ nm, and (d) $b = 65$ nm, respectively.

We now see that the positions of the ridges are represented in the near-field image of the transmission mode by two peaks at the same position. In all cases, the behavior of

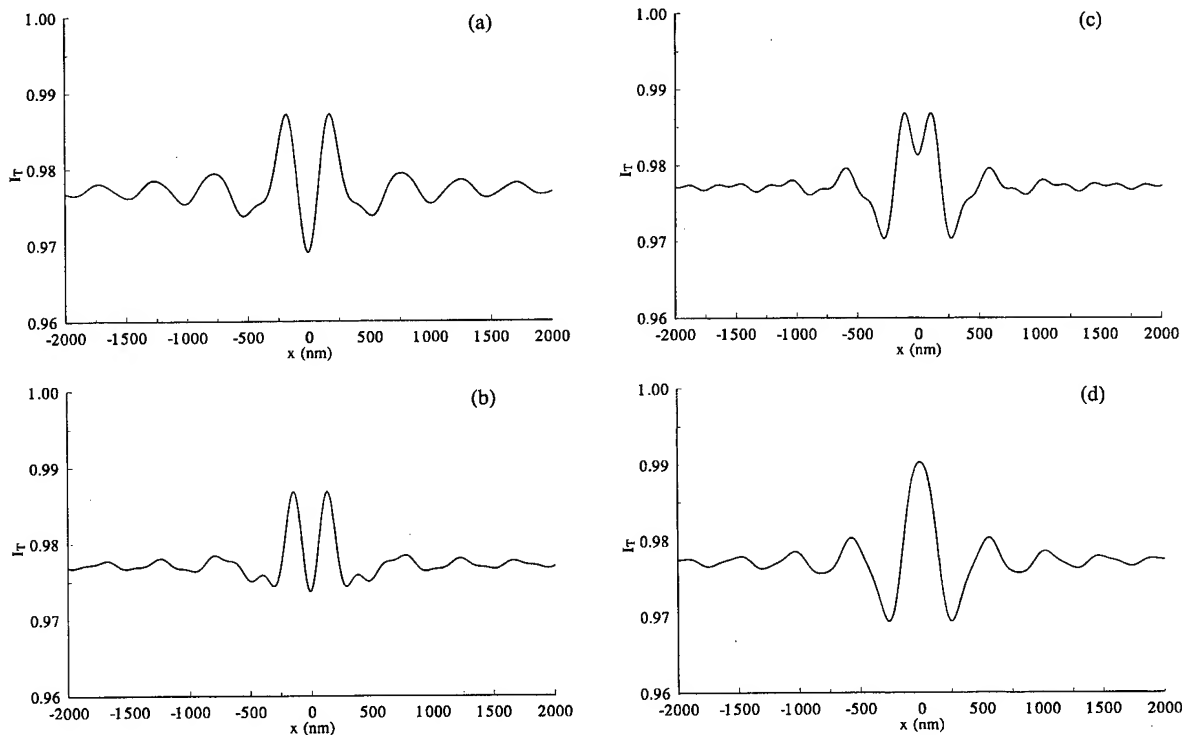


FIG. 4. Normalized near-field intensity I_T in transmission mode, scanning at a constant height above the dielectric surface with two ridges. The probe width and height are $a_g = 50$ nm and $h' = 30$ nm. The ridge width and height are $a_f = 80$ nm and $h = 20$ nm. The separation of the ridges is: (a) $b = 260$ nm, (b) $b = 195$ nm, (c) $b = 130$ nm, and (d) $b = 65$ nm.

the optical image is similar to the case when the objects are grooves: for the case where $b = 260$ nm [see Fig. 4(a)], the separation of peaks is clearly delineated on the image. When the two ridges approach each other, the features of the image get closer [see Figs. 4(b) and 4(c)]. The ridge-peak correlation is lost for $b = 65$ nm, showing a strong peak at the symmetry center of the near-field image [see Fig. 4(d)]. By comparing Figs. 2 and 4 we have observed that the amplitude of the peaks is smaller than that of the dips. This is due to the fact that mean distances of the scatterers (tip and surface objects) are different in both configurations.

In conclusion, we have presented numerical results based on the macroscopic theory¹⁰ for PSTM. The transmitted near-field image through a dielectric substrate with subwavelength objects is illustrated. The feasibility of our approach which enables us to handle surface features of arbitrary shapes and sizes is numerically demonstrated. We have shown that for p -polarized light at normal incidence, the near-field intensity is correlated with the surface structures. The resolution of these features is dependent on the separation of the objects and the tip-object configuration.

ACKNOWLEDGMENTS

The authors acknowledge the financial support by Dirección General de Asuntos del Personal Académico (UNAM, México) under Grant No. IN108995 and CoNaCyT Grant No. 1887 p-E9507.

¹E. Betzig and J. K. Trautman, *Science* **257**, 189 (1992).

²*Near Field Optics*, edited by D. W. Pohl and D. Courjon (Kluwer, Dordrecht, 1993).

³C. Girard and A. Dereux, *Phys. Rev. B* **49**, 11344 (1994).

⁴O. Keller, M. Xiao, and S. Bozhevolnyi, *Surf. Sci.* **280**, 217 (1993).

⁵L. Novotny, D. W. Pohl, and P. Regli, *J. Opt. Soc. Am.* **11**, 1768 (1994), and references therein.

⁶A. Madrazo and M. Nieto-Vesperinas, *J. Opt. Soc. Am. A* **13**, 785 (1996).

⁷J. J. Greffet, A. Sentenac, and R. Carminati, *Opt. Commun.* **116**, 20 (1995).

⁸See, for example, D. Barchiesi and D. Van Labeke, *Ultramicroscopy* **57**, 196 (1995).

⁹D. Maystre, in *Progress in Optics*, edited by E. Wolf (North-Holland, Amsterdam, 1984), Vol. 22, p. 1.

¹⁰S. Wang, M. Xiao, and J. Siqueiros, *J. Mod. Opt.* **43**, 2119 (1996).

¹¹A. A. Maradudin, A. Mendoza-Suárez, E. R. Méndez, and M. Nieto-Vesperinas (unpublished).

Analysis of frictional-force image patterns of a graphite surface

Naruo Sasaki^{a)} and Masaru Tsukada

Department of Physics, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan

Satoru Fujisawa

Mechanical Engineering Laboratory, Namiki 1-2, Tsukuba, Ibaragi 305, Japan

Yasuhiro Sugawara and Seizo Morita

Department of Electronic Engineering, Faculty of Engineering, Osaka University, 2-1 Yamadaoka, Suita-shi, Osaka 565, Japan

Katsuyoshi Kobayashi

Department of Physics, Faculty of Science, Ochanomizu University, 2-1-1 Ohtsuka, Bunkyo-ku, Tokyo 112, Japan

(Received 8 September 1996; accepted 15 February 1997)

We discuss the mechanism of image patterns of the frictional-force microscopy (FFM) of a graphite surface by using a three-dimensional model comprised of a tip connected to a cantilever and a substrate surface. A simulated FFM image is in good agreement with an experimental one. A stable domain of the tip atom position can be defined in an analytic way. In the frictional-force regime, more than one quasistable tip atom position are mapped into a single cantilever basal position. Part of the boundary of the two-dimensional domain of the cantilever basal position appears as a fringe between the bright and the dark areas along the scan direction of the FFM image. General features of FFM images can be completely understood by this analysis. © 1997 American Vacuum Society. [S0734-211X(97)05204-9]

I. INTRODUCTION

The recent development of frictional-force microscopy (FFM)¹ has enabled us to observe features of friction on an atomic scale. Based on the interpretation of FFM images, it has been claimed that the two-dimensional stick-slip motion of the tip atom on the surface provides an important mechanism of atomic-scale friction.²⁻⁵ In previous studies, we studied, by numerical simulations, various features of FFM images that depend on the cantilever stiffness, scan direction, anisotropy of the cantilever, and surface deformation.^{6,7} From the theoretical point of view, however, no detailed analysis of FFM images has at yet been given.

Figure 1(a) shows a FFM image of F_x/k_x obtained by numerical simulation for the case of the constant-height and repulsive force modes. Here we use a three-dimensional model where a single-atom tip connected to a cantilever is scanned on a rigid monolayer graphite surface as shown in Fig. 2(a). The temperature is assumed to be absolute zero, and the tip scanning velocity is almost in the limit of zero in the time scale of lattice vibration. Further, it is assumed that the tip atom is always located at a quasiequilibrium position of the potential surface for each scan position, i.e., the basal position of the cantilever, based on Tomlinson's mechanism.⁸ A raster scan with the x scan repeated in the y direction is performed as shown in Figs. 2(a) and 2(b). The coordinate of the actual tip atom position is denoted as (x, y, z) , and that of the cantilever basal position is denoted as $(x_s, y_s, z_s + l_0)$ hereafter. Here (x_s, y_s, z_s) is the equilibrium

position of the tip atom for the system without interaction with the surface, and l_0 is the natural spring length in the z direction. For each lateral scan position (x_s, y_s) with z_s fixed, the total energy V is minimized based on the Polak-Ribiere-type conjugate gradient method.⁹ The method of simulation is described in detail in Refs. 6 and 7.

Figure 1(b) represents an experimental FFM image. Details of the experimental setup are described in Ref. 5. In order to obtain this experimental image, a raster scan is performed under the constant-height and repulsive force modes. The simulated image of Fig. 1(a) reproduces the experimental one of Fig. 1(b) fairly well. Although the patterns of Figs. 1(a) and 1(b) are different from the shape of the graphite lattice itself, they have the same periodicity as the graphite lattice.

The aim of this article is to clarify the physical meaning of these patterns that appear in the simulated and experimental FFM images by using an analytical method. Based on an analysis of the stable domain of the tip atom position, the "sticking domain" of the tip atom is clearly defined. As will be shown in the following text, a mapping relation between the tip atom position (x, y) and the cantilever basal position (x_s, y_s) , provides an important clue to understanding the FFM image. Domains in the (x_s, y_s) space, which are maps of separate sticking domains in the (x, y) space, are the essential elements of the FFM image. The fringe between the dark and the bright areas in the FFM image along the scan direction is formed by the fringe of the domain in the (x_s, y_s) space.

^{a)}Corresponding author; Electronic mail: naruo@cms.phys.s.u-tokyo.ac.jp

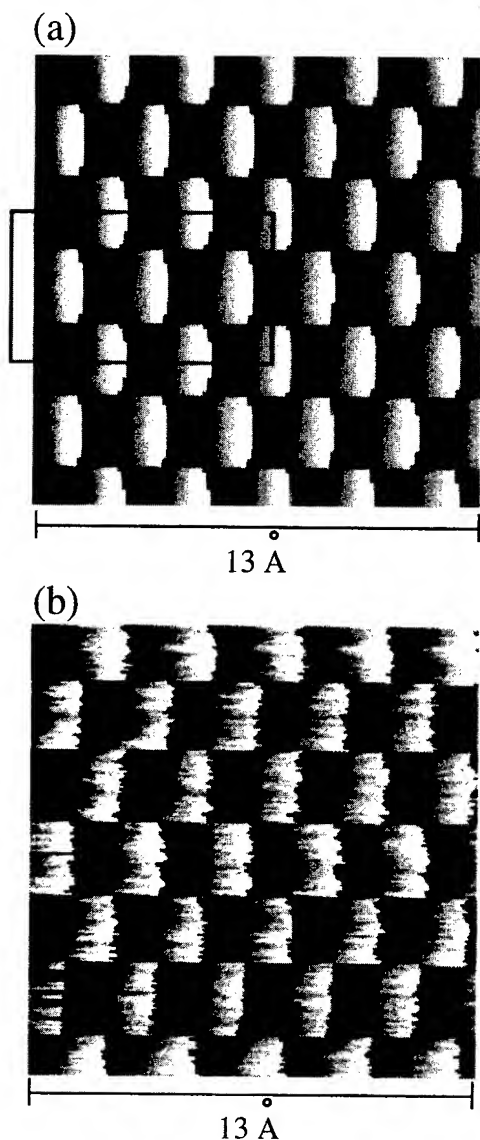


FIG. 1. FFM images of F_x/k_x obtained by (a) numerical simulation and (b) experiment with a tip scanned in the x direction.

II. MODEL POTENTIAL

The total potential energy V_{tot} is assumed to be the sum of the elastic energy of the cantilever V_T and the microscopic tip-surface interaction energy V_{TS} . V_T is assumed to be harmonic, and is expressed as follows:

$$V_T(x, y, z; x_s, y_s, z_s) = \frac{1}{2}(k_x(x - x_s)^2 + k_y(y - y_s)^2 + k_z(z - z_s)^2). \quad (1)$$

The vertical component k_z of the cantilever spring constant is assumed to be 0.25 N/m, which is of the same order as that realized experimentally. On the other hand, the lateral components k_x and k_y are assumed to be parameters. We consider the isotropic spring in x - y plane to be $k_x = k_y = 2.5$ N/m. The cantilever lever basal position is slowly changed from the outside, corresponding to the scan.

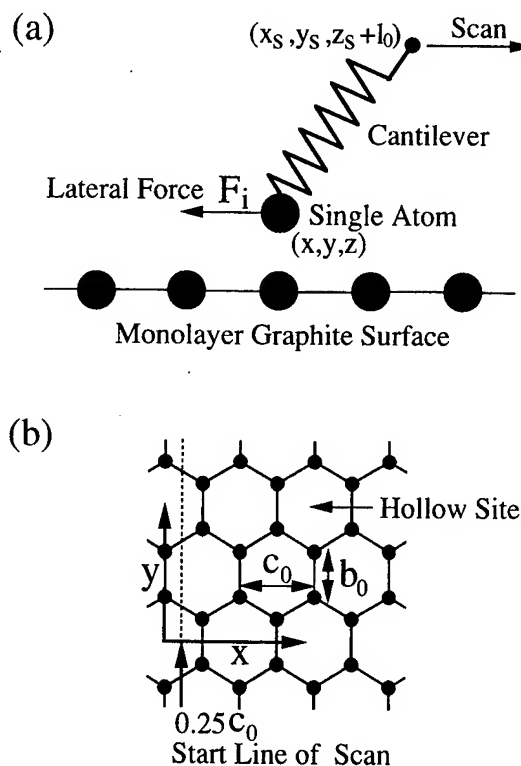


FIG. 2. (a) Schema of the system used in our calculation. The single-atom tip connected to the cantilever spring is scanned on a monolayer graphite surface. (b) The lattice structure of the graphite. c_0 is the length of a unit translational vector of the graphite. b_0 is the C-C bond length of the graphite. $c_0 = 2.46$ Å and $b_0 = 1.42$ Å. The starting line of the scan ($x = 0.25c_0$) is shown by dotted lines.

Tip-surface interaction energy V_{TS} can be obtained as the sum of the pairwise Lennard-Jones potential energies between the single-atom tip and the substrate surface atoms as follows:

$$V_{\text{TS}} = \sum_i 4\epsilon \left[\left(\frac{\sigma}{r_{0i}} \right)^{12} - \left(\frac{\sigma}{r_{0i}} \right)^6 \right]. \quad (2)$$

Here, r_{0i} is the distance between the tip atom and the i th atom in the graphite layer, and the parameters are assumed to be $\epsilon = 0.87381 \times 10^{-2}$ eV and $\sigma = 2.4945$ Å. The interaction potential with these parameters can reproduce the corrugation amplitude of atomic-force microscopy (AFM) images of the graphite quite well.^{10,11}

III. ANALYSIS OF THE FFM IMAGE

We interpret a FFM image by an analytical method whose concept is reported by Gyalog *et al.*⁴ First, the potential V_{tot} is converted from the function of three variables, x , y , and z , into a function of two variables, x and y . For that z should be chosen as the point corresponding to the minimum of V_{tot} , satisfying

$$F_z = - \frac{\partial V_{\text{TS}}(x, y, z)}{\partial z} = k_z(z - z_s). \quad (3)$$

The solution to z of Eq. (3) can be obtained as $z(x,y;z_s)$. V_{tot} is therefore represented as

$$V_{\text{tot}}(x,y;x_s,y_s,z_s) = \frac{1}{2}(k_x(x-x_s)^2 + k_y(y-y_s)^2) + V(x,y;z_s), \quad (4)$$

where

$$V(x,y;z_s) = \frac{1}{2}k_z(z(x,y;z_s) - z_s)^2 + V_{\text{TS}}[x,y,z(x,y;z_s)]. \quad (5)$$

In the next step the lateral coordinates of the tip atom, x and y , can be determined by

$$\nabla_i V_{\text{tot}} = 0 \quad (i=x,y) \quad (6)$$

for each cantilever basal position $(x_s, y_s, z_s + l_0)$.

To satisfy the quasistable equilibrium condition, it is necessary that the Hessian of V_{tot} is positive definite. Therefore the relation

$$|\partial^2 V_{\text{tot}} / \partial x_i \partial x_j| > 0 \quad (i,j=x,y) \quad \text{and} \quad \partial^2 V_{\text{tot}} / \partial x^2 > 0 \quad (7)$$

is required in addition to Eq. (6).

Further, the equilibrium condition, Eq. (6), can be rewritten as follows:

$$(x_s, y_s) = \left(x + \frac{1}{k_x} \frac{\partial V(x,y;z_s)}{\partial x}, y + \frac{1}{k_y} \frac{\partial V(x,y;z_s)}{\partial y} \right). \quad (8)$$

Equation (8) is the condition of the lateral components of the cantilever basal position (x_s, y_s) which can realize the equilibrium tip atom position (x,y) . There are some remarkable features of mapping from (x,y) to (x_s, y_s) . The region of the position (x,y) , where the real solution (x_s, y_s) of Eq. (8) exists, changes with the value of z_s . For larger z_s , the region covers the whole (x,y) space. But with the decrease of z_s , in the frictional-force regime, the region becomes separated into mutually disconnected domains as shown in the top part of Fig. 3. The region of $0 \leq x \leq 3c_0$ and $6.5b_0 \leq y \leq 9.5b_0$ presented in Fig. 3 corresponds to the region surrounded by the frame in Fig. 1(a). Here $c_0 = 2.46 \text{ \AA}$ is the length of a unit translational vector of graphite, and $b_0 = 1.42 \text{ \AA}$ is the C-C bond length of graphite. Hatched domains A-D, A'-C', and A''-C'' in Fig. 3 represent disconnected domains. This type of domain will be called a "sticking domain" hereafter. Each sticking domain is mapped into a domain by Eq. (8) in the (x_s, y_s) space, as shown in Fig. 3. It is assumed that the sticking domains A-D, A'-C', and A''-C'' are mapped into domains A_s -D_s, A'_s -C'_s, and A''_s -C''_s, respectively, and will be called that hereafter. In the frictional-force regime, although the function defining the mapping $[x_s(x,y), y_s(x,y)]$ is univalent, its inverse relation $[x(x_s, y_s), y(x_s, y_s)]$ is generally multivalent. Therefore the mapped domains in the (x_s, y_s) space are partly overlapped by some other mapped domains. In Fig. 3, mutually overlapping domains B_s and C_s in the (x_s, y_s) space, which are the maps of sticking domains B and C in the (x,y) space, are clearly shown by hatched regions.

In Fig. 4, analytically predicted fringes of FFM image patterns for the x scan are shown by the thick curves. For the initial scan condition, it is assumed that no lateral deviation

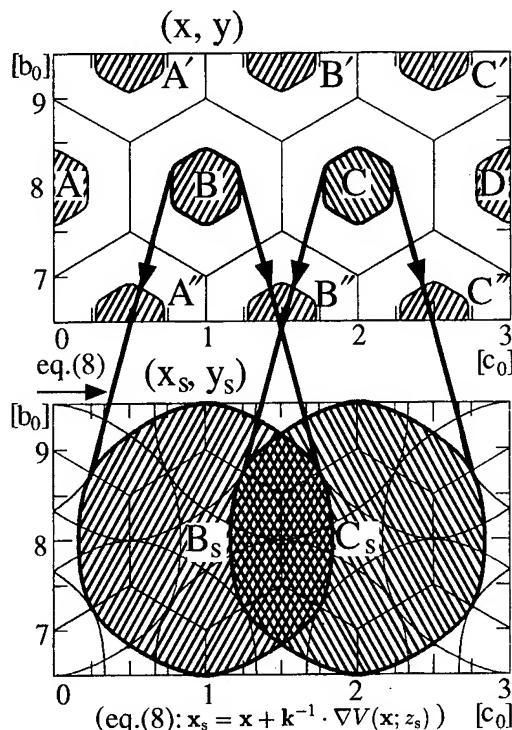


FIG. 3. The top part represents sticking domains A-D, A'-C', and A''-C'' in the (x,y) space indicated by hatched regions. According to Eq. (8), each sticking domain is mapped into a domain in the (x_s, y_s) space, shown in the bottom part. A mapping relation between domains B_s or C_s and sticking domains B or C is clearly seen. Solid lines represent C-C bonds of the graphite surface. The regions of these figures are $0 \leq x \leq 3c_0$ and $6.5b_0 \leq y \leq 9.5b_0$, corresponding to the region surrounded by a frame in Fig. 2(a). Here $c_0 = 2.46 \text{ \AA}$ and $b_0 = 1.42 \text{ \AA}$.

exists between (x,y) and (x_s, y_s) . As shown in Fig. 4, two types of FFM image different in appearance exist in the y direction: region (1) $7.25b_0 \leq y \leq 8.75b_0$, and region (2) $6.5b_0 \leq y \leq 7.25b_0$ or $8.75b_0 \leq y \leq 9.5b_0$. In the case where

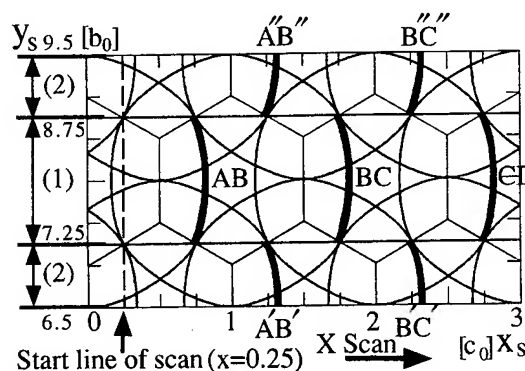


FIG. 4. Thick curves, AB, BC, CD, A'B', B'C', A''B'', and B''C'', represent analytically predicted fringes of FFM image patterns for the x scan. Two types of FFM image different in appearance exist in the y direction: region (1) $7.25b_0 \leq y \leq 8.75b_0$, and region (2) $6.5b_0 \leq y \leq 7.25b_0$ or $8.75b_0 \leq y \leq 9.5b_0$. Solid lines represent C-C bonds of the graphite surface. The starting line of the scan ($x=0.25c_0$) is shown by broken line.

the starting position exists within region (1) in Fig. 4, the tip atom position (x,y) exists within domain A in Fig. 3 for the initial part of the scan. As mentioned above, each (x,y) within domain A is mapped into (x_s,y_s) within domain A_s . Therefore, when the tip atom jumps from domain A to domain B, a thick curve AB, part of the boundary of domain A_s , appears as fringe in the FFM image pattern. For the case of region (2) in Fig. 4, the same explanation as that mentioned for region (1) can be given. Thus it is clarified that not all, but some parts of the boundaries of the domains in the (x_s,y_s) space are observed in the FFM image shown in Fig. 4. Which part of the boundary of the domain in the (x_s,y_s) space is actually observed is determined by the scanning direction. When the load becomes small, or z_s becomes large, the shape of domain (x_s,y_s) changes, and the fringe of the FFM image exhibits a zigzag shape.

IV. CONCLUSIONS

In this work, based on an analytical method, general features of the FFM images of graphite surface can be completely understood. There exists a mapping relation between the tip atom position (x,y) and the cantilever basal position (x_s,y_s) . In the frictional-force regime, disconnected domains in the (x,y) space can be defined as sticking domains. It is clarified that part of the boundary of domain in the (x_s,y_s) space into which the domain in the (x,y) space is mapped appears as a fringe of the FFM image between the bright and the dark areas along the scan direction. This analysis gives a clear explanation to both a simulated and an experimental FFM image.

ACKNOWLEDGMENTS

The authors thank T. Gyalog of the University of Basel in Germany for his very constructive suggestions about interpretations of FFM images. This work was partially supported by a Grant-in-Aid from the Ministry of Education, Science and Culture. The numerical calculations were performed by HITAC S-3800 at the Computer Center of the University of Tokyo.

¹C. M. Mate, G. M. McClelland, R. Erlandsson, and S. Chiang, Phys. Rev. Lett. **59**, 1942 (1987).

²S. Fujisawa, E. Kishi, Y. Sugawara, and S. Morita, Phys. Rev. B **51**, 7849 (1995).

³J. Kerssemakers and J.Th. M. De Hosson, Appl. Phys. Lett. **67**, 347 (1995).

⁴T. Gyalog, M. Bammerlin, R. Luthi, E. Meyer, and H. Thomas, Europhys. Lett. **31**, 269 (1995).

⁵S. Fujisawa, E. Kishi, Y. Sugawara, and S. Morita, Tribol. Lett. **1**, 121 (1995).

⁶N. Sasaki, K. Kobayashi, and M. Tsukada, Phys. Rev. B **54**, 2138 (1996).

⁷N. Sasaki, K. Kobayashi, and M. Tsukada, Jpn. J. Appl. Phys. **1** **35**, 3700 (1996).

⁸G. A. Tomlinson, Philos. Mag. **7**, 905 (1929).

⁹W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, New York, 1992), pp. 413–418.

¹⁰S. A. C. Gould, K. Burke, and P. K. Hansma, Phys. Rev. B **40**, 5363 (1989).

¹¹N. Sasaki and M. Tsukada, Phys. Rev. B **52**, 8471 (1995).

Quantification of topographic structure by scanning probe microscopy

James D. Kiely and Dawn A. Bonnell

Department of Materials Science and Engineering, The University of Pennsylvania, Philadelphia, Pennsylvania 19104

(Received 8 September 1996; accepted 30 April 1997)

Several mathematical approaches for quantifying the three-dimensional topographical structure from scanning probe microscopy images are evaluated. Variational, i.e., scale-dependent, roughness based on root-mean-square roughness, Fourier deconvolution, and the two-dimensional autocovariance function are compared for surfaces with widely varying character in order to develop criteria for accurate quantification. Thermally evaporated gold, a calibration grid, polycrystalline Si_3N_4 , and silicon fracture surfaces serve as models for these techniques. The role of image artifacts on each approach is detailed. © 1997 American Vacuum Society. [S0734-211X(97)13304-2]

I. INTRODUCTION

The topographic structure of surfaces influences a wide range of phenomena, including catalytic activity, wetting behavior, interface formation, and film growth. Furthermore, topographic structure can give quantitative insight into processes that produced a particular surface, for example fracture, cluster aggregation, sputtering, etc. Scanning probe microscopy (SPM) is currently used in a wide range of disciplines to characterize surfaces on scales from μm to nm, often with the objective of quantifying surface textures and shapes.¹ To date, both scanning tunneling microscopy (STM) and atomic force microscopy (AFM) have been used to characterize fracture surfaces of both brittle²⁻⁵ and ductile materials,^{2,6} as well as those from bimaterial interfaces,^{7,8} giving information about energy dissipation during fracture. The topographic structure of surfaces resulting from ion bombardment of graphite,^{9,10} chemical and thermal etching of silicon,¹¹ and metal deposition^{2,12-14} have been related to reaction mechanisms.

The advantage of using SPM for such analyses is that the measurements cover many orders of magnitude of length scale and acquire three-dimensional (3D) data in a digital format. This latter factor is in contrast to most microscopies and allows extensive mathematical treatment of the data. Quantification of surface structure can be based on several fundamental methods of characterizing surface features: using the distribution of heights above or below some reference height, using a spectrum of wavelengths to approximate the surface, or using the correlation of heights over some lateral separation. Calculations commonly applied to SPM data are: root-mean-square roughness,^{15,16} a measure of the variation in height; autocovariance,^{9,16} a measure of spatial correlation of heights; and power density,^{2,13,17,18} a measure of wavelengths of periodic features. These three methods overlap in that they all may be used to calculate the "roughness," but they differ in how they represent the spatial distribution of surface features.

Regardless of mathematical treatment, roughness analyses can be strongly biased by imaging artifacts. The issue of the effect of probe tip geometry on limited spatial resolution and image artifacts has been considered extensively.¹⁹⁻²² It is

widely accepted that the finite size of the tip filters high-frequency components of the topographic structure and introduces inaccuracies in the measurement of high-frequency components. The dependence of average roughness (different than R_q) on tip radius has been calculated,²² but the manner in which more complex tip geometries and other artifacts affect roughness has not. SPM images invariably contain artifacts of the acquisition process that affect roughness calculations. For example, it is unlikely that the sample is exactly perpendicular to the tip; therefore, SPM images usually have some planar artifact (sample tilt) that is not representative of the surface. Roughness measured with this artifact intact is an overestimate, while improper removal of the artifact will also result in misrepresentation. Also, minor temperature fluctuations can cause significant drift in vertical as well as lateral measurements. Additional artifacts can be introduced by optical interference in the detection system, nonlinearity in the x - y motion of the piezoelectric scanner, and random tip jumps.

In order to exploit the utility of SPM in quantifying surface topography, the relative effectiveness of different mathematical approaches must be quantified. The objective of this study was to investigate how roughness measured from SPM images is dependent upon experimental parameters and to evaluate the effectiveness of the different calculations in characterizing topographic structure. After a description of experimental details, three of the most commonly used roughness parameters will be defined and AFM and STM images of several surfaces will be analyzed. The effects of image processing on structure quantification will be demonstrated and the results of the analyses discussed in terms of artifactual content of the results. Finally, the relative merits of three common roughness analyses will be demonstrated on model surfaces and criteria for valid quantification developed.

A. Quantification of topographic structure

The power spectral density (PSD) or the power spectrum, $W(p, q)$, identifies wavelengths of features that contribute to the surface structure. It is the Fourier decomposition of an image into spatial frequencies and is calculated from the 3D image as

$$W(p, q) = \frac{1}{A} \left| \int \int e^{i(px+qy)} h(x, y) dx dy \right|^2,$$

where p and q are lateral frequencies.²³ R_q is related to the power spectrum as follows:

$$R_q = \left[\int_{p_1}^{p_2} \int_{q_1}^{q_2} W(p, q) dp dq \right]^{1/2}.$$

Extraction of the height variation from the power spectra is not direct and requires the assumption that the surface is accurately represented by a series of sinusoids. The PSD is commonly used in the self-affine and fractal descriptions of surfaces.^{2,13,17,18,24}

The two-dimensional (2D) autocovariance function (ACVF), $G(t)$, is a function of distance and direction. Although it is best understood as the expected value of the height product of all two-point pairs that are separated by a length, τ , in some direction, it may be calculated several ways. Defined for a single trace,

$$G(\tau) = \frac{1}{L} \int_0^L h(r)h(r+\tau)dr.$$

The 2D ACVF is calculated by taking the inverse Fourier transform of the product of the Fourier transform and the complex conjugate.²⁵ Thus, the ACVF is closely related to the PSD. The shape of the ACVF may be used to determine some set of correlation lengths in an image,^{16,25} including atomic spacing.⁹ Although the length at which the ACVF first crosses zero is a commonly used correlation length, extracting an unambiguous correlation length from the ACVF is not always straightforward. R_q for the entire image may be obtained from the height of the central, main peak in the 2D ACVF, but since it is obtained using the PSD, it is limited by the sinusoidal approximation.

The advantage of R_q is not only simplicity, but also statistical significance. Since it is the standard deviation of the height, it describes the spread of the height distribution about the mean value. For surfaces that exhibit a Gaussian height distribution, R_q is the difference between the 95th and 5th percentile divided by 3.29.²⁵ Because many surfaces are produced by processes with some degree of spatial randomness (e.g., fracture, chemical vapor deposition, etching), they are expected to exhibit Gaussian or near-Gaussian height distributions, suggesting that R_q is an appropriate descriptor of "roughness." Roughness evaluated from an area, A , may be computed using one of several definitions of R_q . For a three-dimensional $N \times N$ image, where pixel heights are $h(x, y)$, a discrete approximation of R_q is

$$R_q = \sqrt{\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (h_{ij} - \bar{h})^2}.$$

The most apparent limitation of R_q as a structural parameter is the lack of spatial information. A single roughness value provides no insight to the width or spacing of surface features, and roughnesses for surfaces with different spatial variations of features may be identical.²⁶

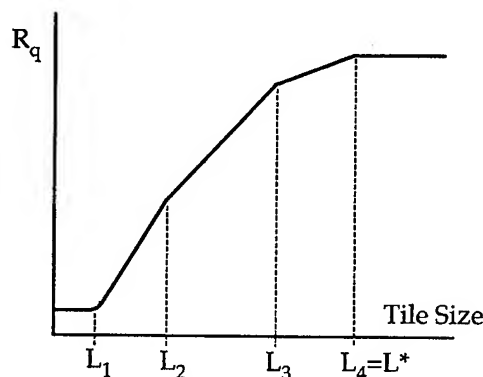


FIG. 1. A schematic of R_q dependence on measurement length. The surface contains four characteristic lengths.

Of course, any measure of surface roughness is scale dependent. As illustrated in Fig. 1, R_q increases with the length over which it is measured, but the dependence may change with tile size. Lengths at which this dependence changes are called characteristic lengths and are here denoted L_1 , L_2 , L_3 , etc. These lengths relate roughness to surface structure and, when roughness is reported in reference to one or several of these lengths, provide a more explicit characterization of the surface. The length scale at which R_q becomes scale independent can be identified as the largest characteristic length of the surface, L^* , providing a definition of the maximum roughness. Analogous to L^* , there is a minimum characteristic length, at which R_q becomes scale independent at small scales, which occurs at atomic dimensions or at a fraction of the tip radius. Between these two extremes, the length dependence of R_q is a function of the distribution of feature heights and widths, and this distribution is described by the slope of the variational R_q . Slope changes identify feature dimensions.

Calculations of R_q on a simple, simulated surface demonstrate this property. A surface often may be viewed as a superposition of several sets of features, each with its own characteristic length and height. R_q is calculated for three simulated surfaces,

$$h(x, y) = 5 \cos(x) \cos(y),$$

$$h(x, y) = 5 \cos(x/5) \cos(y/5),$$

$$h(x, y) = 5 \cos(x) \cos(y) + 5 \cos(x/5) \cos(y/5).$$

A line profile from the third surface appears in the inset of Fig. 2(a), along with variational roughness from the three surfaces. The characteristic lengths of the first two surfaces, L_1 and L_2 , respectively, are identified in Figs. 2(a) and 2(b) from curves A and B, which present variational R_q from the first two surfaces. These characteristic lengths are approximately 2/3 of the period of oscillation of each surface. Variational R_q from the superposed surface, curve C, identifies each of the characteristic lengths and quantifies the amplitudes of superposed features, i.e., R_q associated with the short-wavelength characteristic length is that of one set of features ($R_q = 2.5$), while that of the long-wavelength features quantifies the superposition of the wave forms. (The

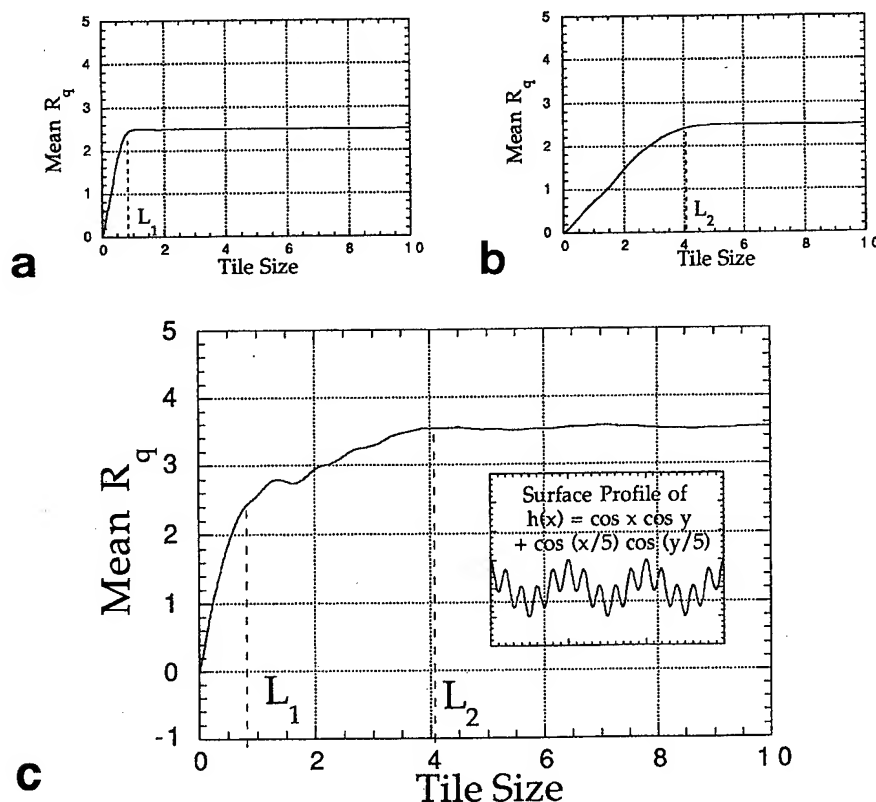


FIG. 2. Calculated mean variational R_q for three simulated surfaces. (a) $h(x,y) = 5 \cos(ax)\cos(ay)$, (b) $h(x,y) = 5 \cos(x/5)\cos(y/5)$, and (c) $h(x,y) = 5 \cos(x)\cos(y) + 5 \cos(x/5)\cos(y/5)$. A single characteristic length is identified in (a) and (b), while two are identified in (c). The two characteristic lengths in (c) match those in (a) and (b).

units are arbitrary in this analysis and the minimum information density is 100 points per high-frequency period.) From this example, it is clear that variational R_q identifies multiple feature lengths.

In the scale-dependent regime, R_q often follows a power-law dependence on L , and is often written $R_q \propto L^\alpha$.¹² The exponent α may vary between 0 and 1, and depends on the local surface structure. When α is noninteger and constant over some range of length scales, the surface may be described as self-affine, and may be quantified using fractal descriptions. The exponent α is related to the local fractal dimension, D , through the relation $\alpha = 3 - D$.¹³ This approach to analysis of SPM data has been recently reviewed.

II. EXPERIMENTAL PROCEDURE

Three types of surfaces were evaluated:

- (1) Surfaces with features of a single width: A Au-coated calibration grid with 10 μm periodicities in the x and y directions and a depth of 200–210 nm was imaged using AFM.
- (2) Surfaces with features of a narrow range of widths: The flat face of the calibration grid and Au-coated Si_3N_4 were imaged using STM. Individual Si_3N_4 grains ranged from widths of 50 to 180 nm, but agglomerations of these grains were as wide as 1 μm . The features of the calibration grid face ranged between 20 and 50 nm.

- (3) Surfaces with features of a range of widths spanning from 10 to 10^5 nm: Single-crystal Si fracture surfaces were imaged using AFM. The $\{110\}$ surface had been oriented as the tensile surface and the fracture plane was $\{110\}$. These fracture surfaces are well studied and feature widths range over several orders of magnitude. Because of this wide range and the self-affine structure of these features, these surfaces are often characterized using fractal descriptions.^{2,5} The transition between the mist and hackle regions was characterized; only one image and analysis, thereof, is presented here.

AFM images were acquired (Digital Instruments Dimension 3000) in contact ac mode using a long-range scanner (90 μm lateral travel, 6 μm vertical travel) in ambient air. Cantilevers were single-crystal silicon with spring constants in the range of 25–73 N/m. The operating set point was approximately 70% of the free-oscillation amplitude and scan rates were in the range of 0.1–0.5 Hz. STM images were acquired (Digital Instruments Nanoscope III) in ambient air in constant-current mode using a short-range scanner (9.5 μm lateral travel, 3 μm vertical travel) using a Pt–Ir tip. The operating bias was 600 mV, the current set point was 0.6 nA, and the scan rate was 0.5 Hz. AFM and STM images consisted of 256×256 data points.

In order to determine the effect of image processing on calculated roughness, images of the calibration grid were analyzed after each step of a typical postacquisition modifi-

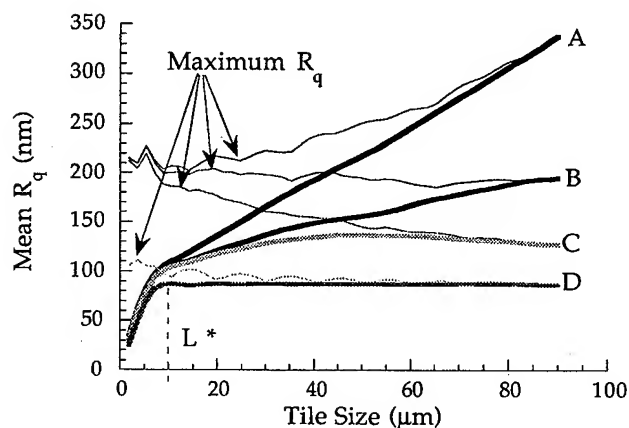


FIG. 3. The effect of postacquisition image processing on calculated root-mean-square roughness for a $90\text{ }\mu\text{m}$ section of a calibration grid. The mean R_q for all tiles of a given size in one image is represented by the bold line, and the maximum R_q by the thin line. (A) raw data, (B) plane-fitted in the fast-scan direction, (C) plane-fitted in both fast- and slow-scan directions, and (D) plane-fitted and flattened. The characteristic length of the grid, $10\text{ }\mu\text{m}$, is identified from these curves.

cation procedure. That procedure was: (1) remove a plane in the x direction that had been least squares fitted to the average of all rows of data (plane-fitting), (2) do the same in the y direction, and (3) raise or lower each line of the image until the mean average of each line was at the same height (flattening). Subsequent images were analyzed after the full procedure was completed. Plane fitting and flattening were linear, i.e., no higher-order fits were used. Effects of tip convolution on roughness were determined by imaging similar areas of the Si_3N_4 using a single and a double STM tip.

The following algorithm was used to calculate variational roughness over a range of lengths:

- (1) A square "tile" was defined to be a subset of the image with dimensions of five pixels by five pixels. Roughness was calculated for each tile (for a five-pixel tile size in a 256×256 data file, $251 \times 251 = 63\,001$ values were calculated). The mean of values for this tile size was calculated and recorded as the roughness for that length.
- (2) The tile size (L) was increased by increments of five pixels ($10\text{ pixels} \times 10\text{ pixels}$, then $15\text{ pixels} \times 15\text{ pixels}$, etc.), and the mean of each tile size recorded. The largest tile size was 255×255 , approximately equal to the image size ($L = L_i$).

The variational root-mean-square roughness (R_q), power spectrum (PSD), and autocovariance function (ACVF), were calculated for each image.

III. RESULTS

The effect of plane-fitting can be most easily quantified in images of a calibration grid. Calculations on the raw data and after each image processing procedure reveal that both plane-fitting and flattening reduce mean R_q values, but the certain character of the variational R_q remains intact. This is illustrated in Fig. 3, which compares the scale-dependent R_q calculated for the raw data, data plane-fitted in the x direction

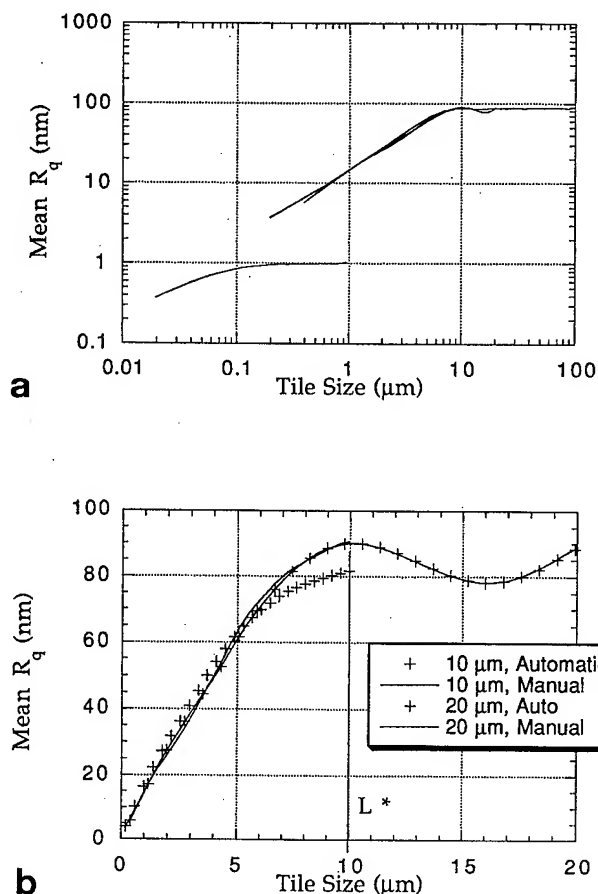


FIG. 4. The effect of image size on R_q for images of the calibration grid. (a) Mean R_q calculations from 1, 10, 20, and $90\text{ }\mu\text{m}$ images, which had only the sample tilt removed. The 10, 20, and $90\text{ }\mu\text{m}$ data coincide, while the $1\text{ }\mu\text{m}$ data represent the roughness of the fine structure on top of the grid. (b) Mean variational R_q for 10 and $20\text{ }\mu\text{m}$ images, which had only the sample tilt removed (labeled "Manual") or had the least-squares-fit plane removed (labeled "Automatic"). Automatic plane-fitting reduces R_q when the image size is smaller than or equal to L^* (cf. $10\text{ }\mu\text{m}$ curves), but not when the image size is greater than L^* . The two $20\text{ }\mu\text{m}$ curves are coincident.

only, data plane-fitted in both x and y directions, and data plane-fitted in both directions and flattened. Along with the mean R_q (bold lines), the maximum R_q is presented (thin lines) to indicate the scatter in the measurements. Regardless of the postacquisition image processing, the characteristic feature length, L^* , of the grid is correctly identified by a slope change at $L = 10\text{ }\mu\text{m}$. L^* cannot be identified from a single value, as illustrated by the maximum R_q , which exhibits no characteristic length. Plane-fitting reduces R_q , but when measured at lengths greater than L^* , the amount of reduction depends on the degree of image processing and varies between being indeterminate (A) and 70 nm (D). It is significant that the effect of plane-fitting on R_q is strong and increases linearly with tile size.

Since plane subtraction derives the plane from the entire image, changing the image size will affect the calculated roughness. A comparison of calculations from different-size images is shown in Fig. 4(a), which presents data from 1, 10, 20, and $90\text{ }\mu\text{m}$ images. The $1\text{ }\mu\text{m}$ curve was taken from the flat top of the grid, so it quantifies only the fine structure.

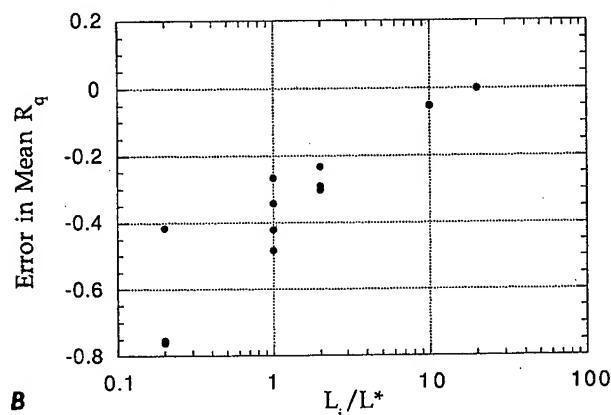
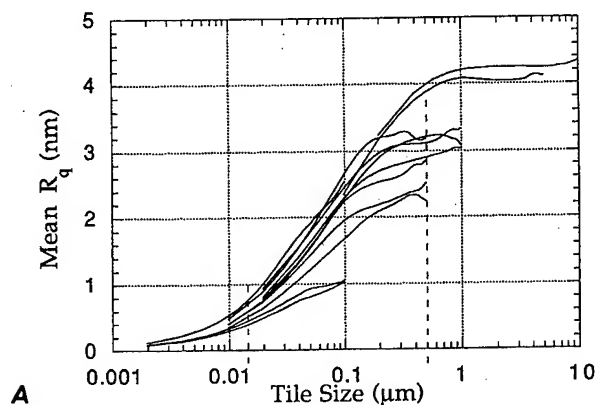


FIG. 5. The effect of image size on R_q for images of gold-coated Si_3N_4 . (a) Mean variational R_q for 0.1, 0.5, 1, 5, and 10 μm images. When the image size is smaller than the characteristic length, improper plane fits reduce R_q at moderate and larger L/L_i , but not as significantly at lower L/L_i . The characteristic length for the surface is approximately 1 μm . (b) The error associated with R_q measured at $L=L_i$ as a function of L_i/L^* . Reducing the size of the image increases both the magnitude of the error and the variability in results.

The 10, 20, and 90 μm variational R_q quantify the structure of the grid and are nearly coincident except at larger tile sizes, where the 20 μm curve deviates from the others. This is believed to be due to averaging effects involved with the changing proportion of the image area that can be described as either high or low areas of the grid. Aside from this, the three curves demonstrate that when only the sample tilt is removed, different-size variational R_q of the same structure have the same form.

The effect of an improper plane subtraction on roughness is demonstrated in Fig. 4(b), which compares 10 and 20 μm images that have been automatically plane-fitted with those that have had only the sample tilt subtracted. As a point of clarification, proper plane fits remove only the sample tilt artifact and improper plane fits remove some of the surface structure when the best-fit plane is removed. Automatic plane-fits are often, but not always, improper. The effect of an improper automatic plane fit is observed in the 10 μm image and not in the 20 μm image, suggesting that for a sufficiently

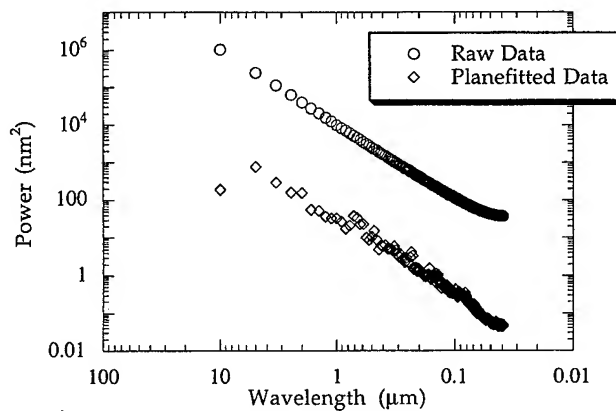


FIG. 6. The effect of plane-fitting on PSD of the image of a silicon fracture surface.

large image size, the automatic regression correctly identifies the sample tilt. The error involved in improper plane fits is largest when the image size, L_i , is smaller than L^* , as is demonstrated in calculations from the Si_3N_4 surface. Figure 5(a) presents roughness from 100 nm, 500 nm, 1 μm , 5 μm , and 10 μm size images of the Si_3N_4 surfaces, which were automatically plane-fitted and flattened. The curves overlap when L/L_i is small and, if the region in which $L/L_i > 0.25$ in each curve is ignored, a variational R_q curve similar to that of Fig. 1 may be constructed. This composite variational R_q quantifies the true roughness of the surface, since it adjoins measurements in which plane-fitting errors are minimized and identifies two characteristic lengths; one at 20–30 nm and one at 400–600 nm. Note that the slope changes occur over a range of lengths and that the range is a function of the distribution of feature sizes. (This is particularly true for the small length, which must be approaching that of the tip.) R_q 's measured at $L=L_i$ are noncoincident, suggesting that the accuracy of R_q measurements was compromised at larger L/L_i by plane-fitting. All images smaller than L^* (about 500 nm) have R_q values less than those calculated at lengths greater than L^* , suggesting L^* as a criteria for maximum R_q measurement. If we accept R_q measured at $L=10 \mu\text{m}$ as valid, we can determine the relative error in the roughness measured as a function of image size. Figure 5(b) presents this error as

$$\text{error} = \frac{R_q|_{L=L_i} - R_q|_{L=10 \mu\text{m}}}{R_q|_{L=10 \mu\text{m}}}$$

The magnitude of error decreases as L_i/L^* approaches 2, as does the variability of R_q measurements.

Plane-fitting affects not only R_q calculations, but those of PSD and ACVF as well. The effect of plane-fitting on PSD calculations is shown in Fig. 6. PSD calculations were performed on the raw and plane-fitted images of the silicon fracture surface and demonstrate the reduction of the power at all wavelengths. The consequence is to reduce the amplitude at all frequencies similarly, except that associated with the image size, since that is the frequency of a background plane.

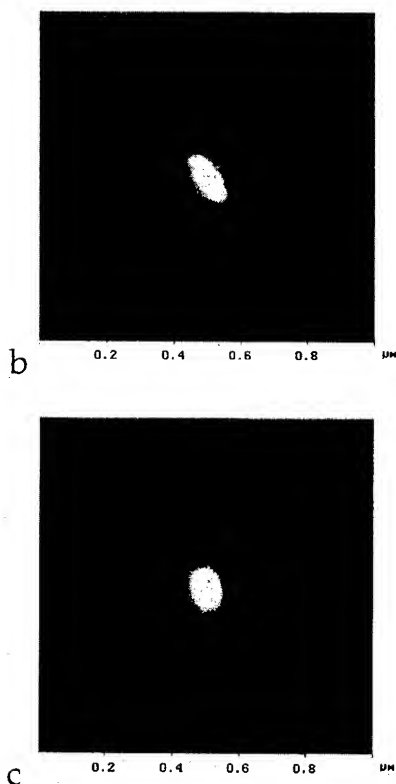
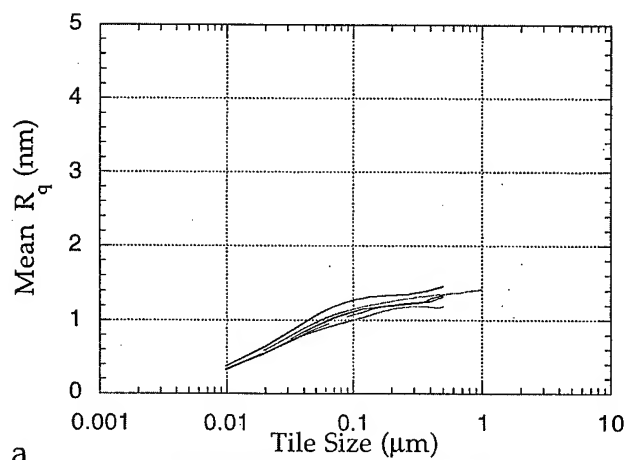


FIG. 7. The effect of tip artifact on R_q and ACVF. (a) R_q measured from a STM image of Si_3N_4 acquired using a double tip. Magnitudes are a factor of 2 lower than those acquired with a good tip [cf. Fig. 5(a)]. (b) ACVF for the image acquired with a double tip. Satellite peaks appear at the separation and direction of the two tips. (c) ACVF acquired from the same area using a good tip. No satellite peaks are present and the width of the central peak more accurately corresponds to surface features.

Thus, the slopes of the power spectra, which are associated with the dimension of the surface, are the same.

The effect of tip geometry on roughness may be observed in several calculations. Figure 7 shows a series of variational roughness curves acquired from the same area from which the data of Fig. 5(a) was acquired, but with a different tip. Roughness is lower by a factor of 2 and no single L^* is apparent, but no other information can be extracted from the analysis. The 2D ACVFs for the 500 nm images from the

bad and good tips are presented in Figs. 7(b) and 7(c). These plots compare the ACVF as a function of lag length and direction, where the lag length and direction are measured relative to the center of the 2D ACVF. The plots have been normalized such that the vertical range is equal to the height of the main peak so that the widths of features in the ACVFs can be compared. The most notable difference between the two sets of data is the appearance of two satellite peaks in the ACVF from the bad tip. If it is assumed that Si_3N_4 has an isotropic structure that did not change between images (a reasonable assumption), anisotropies in the ACVF are attributable to artifacts, and the presence of a second peak suggests a double tip (two peaks appear because the ACVF is a symmetric construct). In Fig. 7(b), two secondary peaks appear on opposite sides of the main peak, approximately 60 nm from the center of the main peak, the same repeat distance of features in the STM image. The elongation in the central peak in Fig. 7(c) is likely due to thermal drift.

In order to compare the efficacy of various roughness calculations, the variational R_q , power spectrum, and ACVF are calculated for a model surface (the calibration grid) and two real surfaces; one nearly isotropic, one with considerable directionality. Variational R_q from the calibration grid [Fig. 8(b)] exactly identifies the grid periodicity as does the PSD [Fig. 8(c)] as 10 μm . The power spectrum, however, reveals the error associated with Fourier based calculations when a periodic square-wave surface is been approximated by a series of sine waves with varying power. The oscillating power density is an artifact of the approximation. The ACVF, shown in Fig. 8(d), reproduces the periodicities of the surface in all directions and illustrates the regularity and flatness of the grid. The distortion at the edge of the image is associated with nonlinearity in the scanner. This distortion is not readily evident in the image but is clear in the ACVF. Spatial variations calculated from the ACVF and variational R_q are nearly identical.

The results of R_q , power spectrum, and ACVF calculations for a 1 μm area on the face of the grid are shown in Fig. 9 along with the 1 μm image. Variational R_q [Fig. 9(b)] manifests a characteristic length of 50–80 nm. The information density in this image (1 pixel each 4 nm) is not sufficient to identify the smaller characteristic length from this calculation. At a wavelength approximately equal to this length, the power spectrum, [Fig. 9(c)] reaches a maximum, but a single characteristic feature length is not discernible. The radius of the center peak in the ACVF [Fig. 9(d)] is 15–25 nm, which correlates well with the apparent grain size shown in Fig. 9(a). For this surface, the PSD is relatively ineffective, while variational R_q and the ACVF identify distinct characteristic lengths.

Analysis of the single-crystal silicon fracture surface provides a comparison of the merits of the different techniques on a directional, self-affine surface. The 10 μm image [Fig. 10(a)] illustrates the wide range of lengths and strong directionality of fracture surface features. Since some surface features were larger than the vertical travel of the AFM, the largest features are not quantified by any of the techniques.

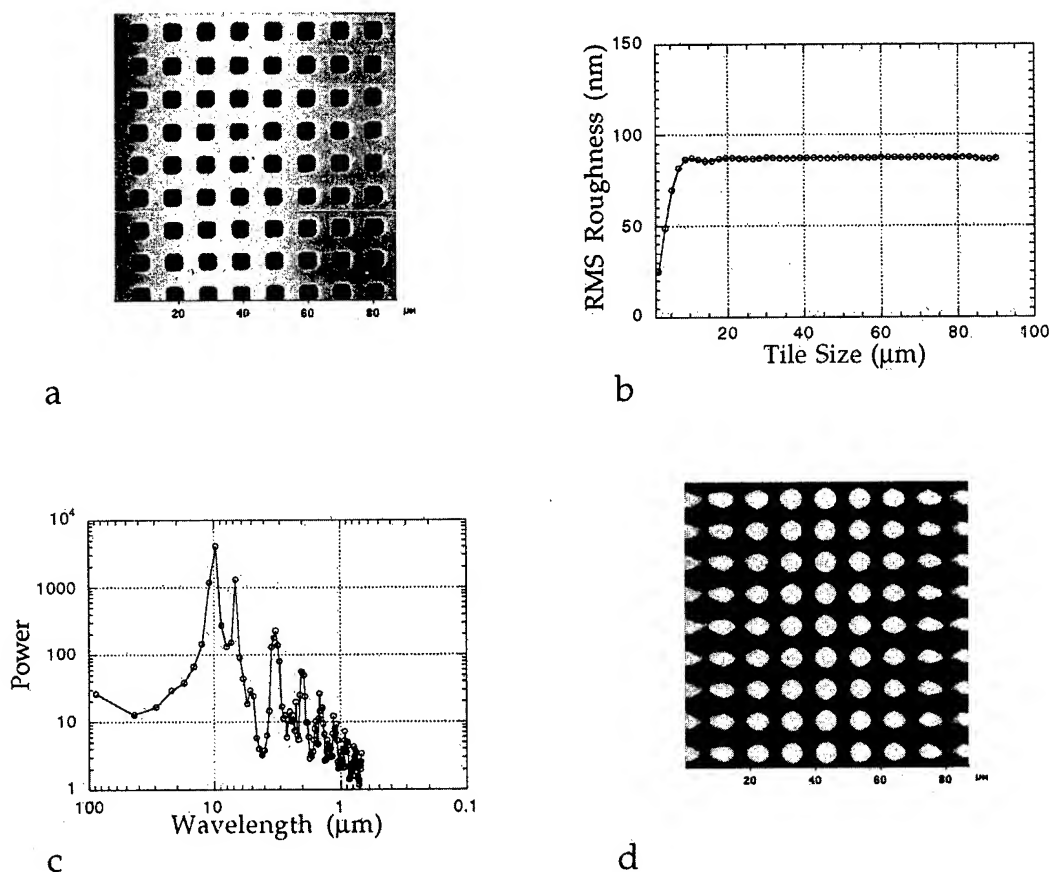


FIG. 8. (a) AFM image of calibration grid. (b) Variational R_q . The characteristic length of $10 \mu\text{m}$ is identified. (c) PSD. (d) 2D ACVF. The ACVF reproduces characteristic lengths, periodicity, and directionality.

The variational R_q [Fig. 10(b)] identifies a characteristic length between 500 nm and $2 \mu\text{m}$, but it is not a characteristic length that identifies the scale-independent regime. No characteristic length greater than this appears, implying R_q is scale dependent over the entire scan range and no single roughness value characterizes the surface at lengths between 200 nm and $10 \mu\text{m}$. The linearity of the PSD, shown in Fig. 10(c), suggests that a single dimension (>2.0) describes the surface, but the power at the largest wavelength deviates from linear behavior. The ACVF [Fig. 10(d)] reproduces the directionality of the surface, but, unlike the R_q and the PSD, it identifies a characteristic feature length, the size of which changes with orientation. The width of the peak ranges between 700 nm and $2 \mu\text{m}$. A further comparison between the methods can be made on the ability to describe the self-affine nature of the surface. While the ACVF offers no such description, the local fractal dimension may be calculated from the slopes of the variational R_q and the PSD. The slope of the PSD gives a dimension of 2.46 , while the slope of the variational R_q gives a dimension of 2.72 . The discrepancy between these numbers arises from the contrasting methods of determining the scaling behavior. Because fracture surfaces are self-affine and not self-similar, a different definition of the local fractal dimension arises from each method.²⁴ Each method quantifies the surface locally using a different technique; it is not expected that these techniques scale ac-

cording to the same law. The approximation of this surface by sine waves (PSD) and the scatter of its height data about the mean (R_q), apparently scale much differently, and, as observed by other researchers,¹³ the R_q technique results in a higher scaling exponent.

The results on all three surfaces are summarized in Table I.

IV. DISCUSSION

These results illustrate both the capability and the pitfalls associated with quantifying 3D surface structure from SPM images and underline the need for specific criteria for evaluating the validity of roughness calculations. The most important issues are the effect of image size, of image modifying operations, and of tip size on calculated roughness.

Criteria for evaluating the validity of calculated roughness at a particular image size are best discussed in terms of variational root-mean-square (rms) values, but are generally applicable. Since R_q quantifies the scatter of the height about the mean, it is best understood by considering the height distributions of the images. A surface has some distribution of heights, which may or may not be Gaussian and, when imaged locally, only a fraction of the surface height distribution appears in the image height distribution. Changing the relative size of the image changes the distribution in the lat-

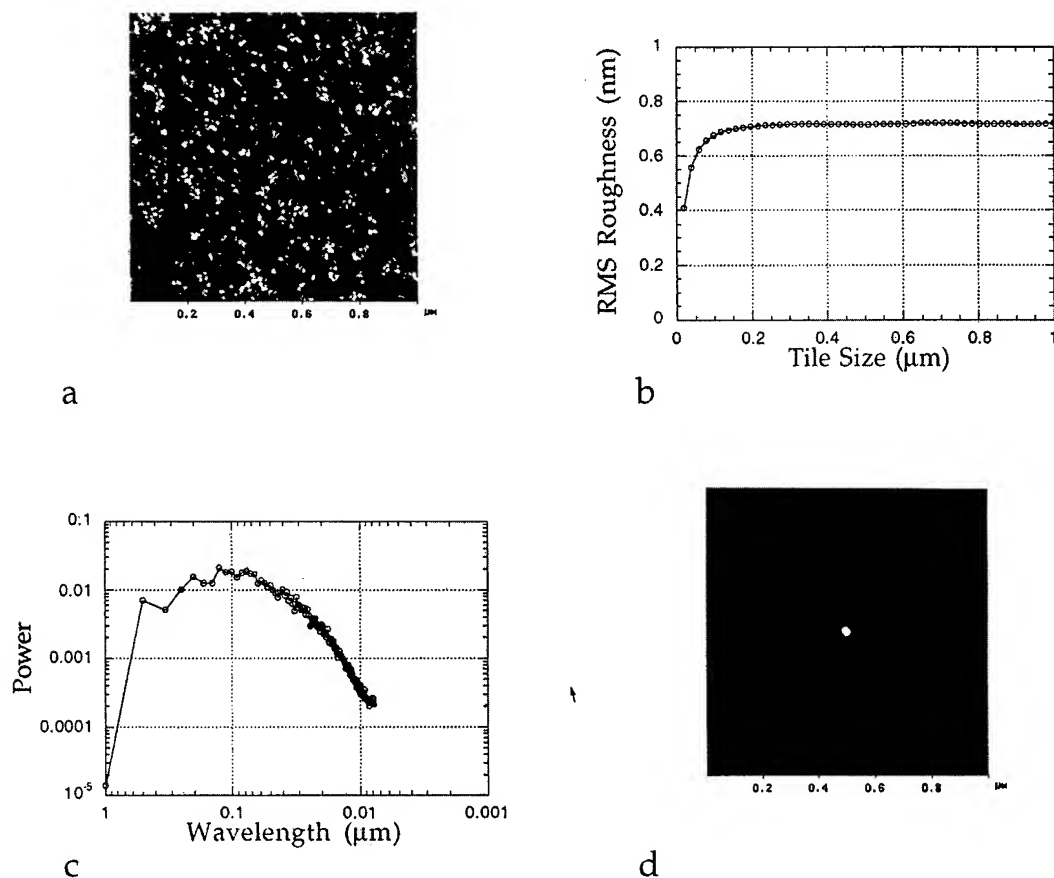


FIG. 9. (a) STM image of the face of the calibration grid. (b) Variational R_q . L^* appears to be 50–80 nm. (c) PSD. The PSD flattens at 70–100 nm, but extraction of a characteristic length is difficult. (d) 2D ACVF. The radius of the peak identifies 15–25 nm as a characteristic length, which correlates well with observed grain sizes.

ter until the characteristic length of the surface is reached, at which point the surface may be described using a single R_q measurement.

For example, an image of a calibration grid [Fig. 9(a)] has area fractions of low areas ("pits") and of high areas, which are separated vertically by 200 nm. The height distribution is, thus, bimodal with one sharp peak centered roughly 200 nm above a second, smaller sharp peak. Changing the image size changes the ratio of high and low areas, which changes the image mean height and the distribution of heights about the mean (R_q). Changes in R_q with image size will become minimal when the area ratio ceases to change significantly with image size. Confidence in a single scale-independent R_q measurement increases with image size and, for the calibration grid R_q measurements, oscillates about 87.5 nm with the amplitude of the oscillations decreasing with image size until it is 0.5% of R_q at 40 μm . For this surface, then, R_q measured at four times the largest feature length is accurate.

The surface structure of the gold-coated Si_3N_4 (Fig. 5) is markedly different, as a single, near-Gaussian peak characterizes the height distribution of the surface. The same argument applies, however. At small measurement lengths, the average height distribution is that of only a few features or of a fraction of a feature, while at large lengths many features are included, including those not apparent to the eye. The

slope changes that identify characteristic lengths for this surface are broader, indicative of a broader feature size distribution, but above approximately 1 μm , R_q is nearly scale independent. The largest characteristic length of the surface appears to be on the order of 400–600 nm, even though to the eye the typical feature size is 200 nm. Apparently, long-range features comprised of many grains exist and it is these features that determine L^* . At lengths greater than 2 μm , slight variations in the height distribution give slight R_q fluctuations, but only by roughly 5%.

This comparison illustrates that the measured roughness becomes scale independent at a length of 4–5 L^* , but also that the determination of L^* is not always straightforward. A measure of the characteristic length from a visual inspection may grossly underestimate the actual characteristic length, as was the case for the Si_3N_4 . The true L^* can only be determined by performing variational R_q measurements over several orders of magnitude.

The most influential image-modifying operation appears to be plane-fitting. R_q from poor plane-fitting operations can be significantly different from that of the true surface, as illustrated in several of the previous figures. Proper sample tilt removal reduces R_q from raw data and does so by an amount that increases with tile size, as is seen in the decreasing slopes of the mean and maximum variational R_q after

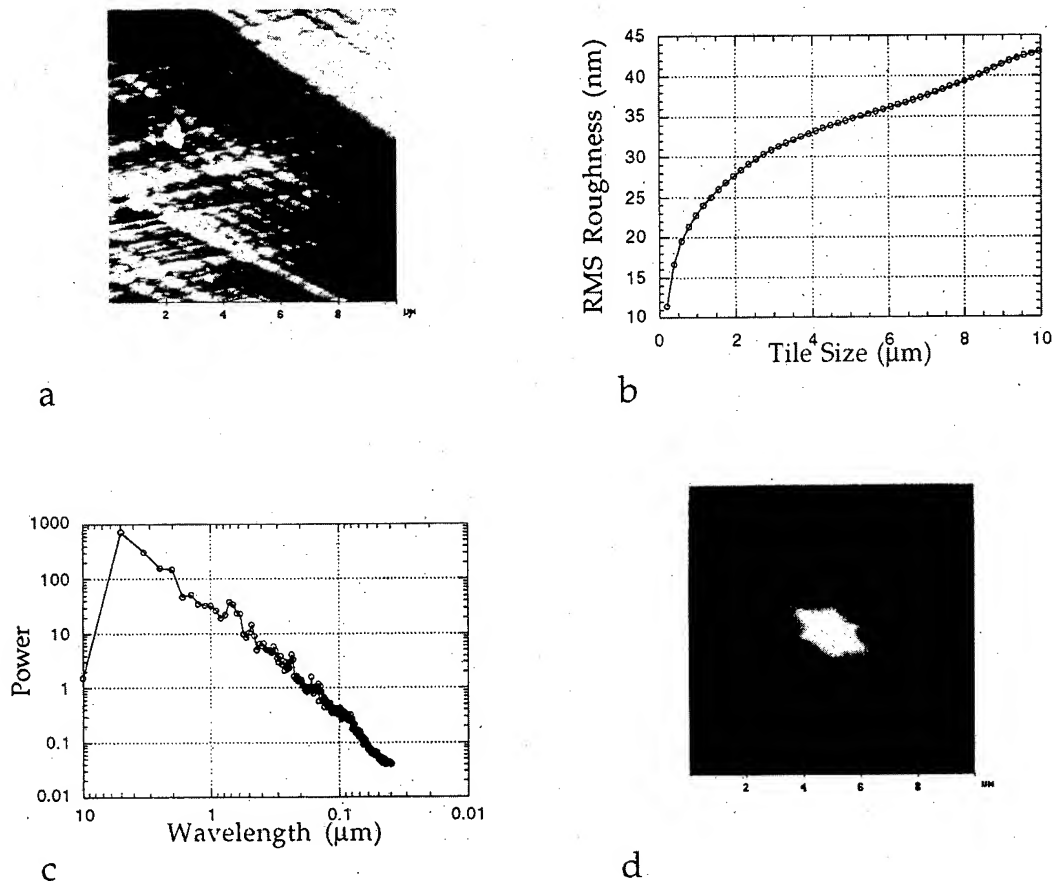


FIG. 10. (a) AFM image of silicon fracture surface at the transition between mist and hackle regions. (b) Variational R_q . It exhibits no upper characteristic length, but an intermediate characteristic length between 500 nm and 2 μm can be identified and correlated to the striations on the surface. (c) PSD. No characteristic length appears in the PSD; the surface may be described as self-affine. (d) 2D ACVF. The ACVF quantifies the directionality of the surface and peak widths vary with direction. The width ranges from 700 nm to 2 μm, suggesting a characteristic length varies with direction.

each plane-fitting operation in Fig. 3 and in the marginal change in R_q for small L . To understand this, the components of measured roughness can be separated into that of the surface and those due to artifacts of the imaging process. For any image (or tile of an image) of size L , the planar artifact due to sample tilt may be written as $h(x,y) = ax + by + c$, and the roughness of this plane alone is given by $R_q^{\text{plane}} = L\sqrt{(a^2 + b^2)}/12$, so increasing the tile size increases the component of R_q due to sample tilt.

While the result of sample tilt removal is a reduction in R_q on all length scales, improper plane-fits can result in cal-

culated roughness that is lesser or greater than the real R_q , depending on L/L_i . Improper plane-fitting reduces R_q when $L = L_i$, but may increase it at other tile sizes. The automatic, improper plane-fit of the 10 μm data in Fig. 4(b) increased R_q over that of the manual, proper plane-fit for tile sizes smaller than approximately 5 μm and reduced it for tile sizes greater than approximately 5 μm. This exaggeration at small L/L_i and reduction at large L/L_i results in characteristic lengths that are artificial and may be considerably smaller than the real values. The 10 μm variational R_q in Fig. 4(b) and the variational R_q from 100 nm, 500 nm, and 1 μm

TABLE I. Results of quantification techniques on three surfaces.

	Grid	Grid, fine structure	Si ₃ N ₄	Silicon
Apparent feature size	10 μm	20–30 nm	200 nm	Many
Characteristic length of variational R_q	10 μm	50–80 nm	400–600 nm	500 nm–2 μm
Characteristic length of PSD	10 μm	> 70 nm	...	None
Characteristic length of ACVF	10 μm	15–25 nm	...	700 nm–2 μm

images in Fig. 5(a) demonstrate this effect. All exhibit apparent L^* s, which differ from that of the true surface, and the R_q measured at $L=L_i$ varies widely, in all cases being smaller than that resulting from a proper planefit.

One goal is to determine a criterion to identify when a planefitting operation correctly removes the sample tilt such that none of the surface structure has been removed along with it. Since the error due to the planefit in the 20 μm mean variational R_q in Fig. 4(b) is negligible, it may be assumed that this automatic planefit was proper. This automatic planefit was a proper planefit when the L_i was only twice the largest feature size. The largest feature size of the Si_3N_4 images appears to be on the order of 200 nm, while a proper planefit occurred only for images more than an order of magnitude greater [Fig. 5(b)]. This suggests that the 200 nm features lie upon a structure with a longer wavelength and that the most obvious topographic structure is not always the largest structure. The eye may not identify less pronounced structures and it may be incorrect to base a roughness measurement on a visual inspection of the SPM image.

One of the greatest concerns in SPM is the degree to which the probe radius affects imaging. Sharp tips reveal accurate feature widths and faithfully probe both peaks and valleys of the surface, while tips with larger radii produce flatter images with wider features. This topic has received considerable attention in terms of strategies for image reconstruction,¹⁹ etc. Here, we discuss only the effect on roughness as calculated from the relationships described above. The finite tip size acts like a low-pass filter removing wavelengths smaller than the tip size. The influence of tip morphology on R_q measurements is dependent upon feature dimensions and the length at which roughness is measured, but for a given surface, R_q decreases with increasing tip radius. This effect is seen by comparing Figs. 5(a) and 7. The R_q obtained with a double tip (Fig. 7) is roughly half that of R_q obtained with a single tip. The "good" tip is believed to have a radius of approximately 10–20 nm, while the double tips were found to have a separation of approximately 60 nm. Since R_q is a function of tip morphology, it is problematic to compare R_q results on scales on the order of the tip radius from images obtained using different tips.

With some assumptions regarding surface shape, tip morphology can be analyzed using the 2D ACVF. Figure 7 demonstrates this for the case of imaging a sample with single and double tips. Double tips produce strong secondary peaks at the lag length and direction corresponding to the separation and orientation of the two tips. The tip radius may be estimated from the primary peak, as demonstrated in Figs. 10(a) and 10(b). The diameter of the primary peak in the ACVF from the AFM image is approximately twice that from the STM image, and since the surface was approximately the same in both, this difference is due solely to tip geometry. This estimate could be used to determine the length below or the frequency above which the data is most affected by tip convolution.

The three mathematical treatments provide spatial descriptions of features, but none is universally applicable.

Each is a unique construct and characterizes a different aspect of the surface, so each is best suited for certain surfaces. Surfaces that can be well approximated by a series of sine waves will be well described by the PSD, and those that cannot result in power spectra that are difficult to interpret. The PSD in Fig. 9(c) shows several "peak" wavelengths, which have higher powers than the surrounding wavelengths and are not directly associated with the structure of the surface. The grid is better described using the ACVF or the variational R_q . Variational R_q identifies the dominant wavelength of the surface, while the ACVF reproduces the wavelength, directionality, and periodicity. A surface typical of many engineered structures, that of the top of the grid, is more amenable to the PSD, so that all three methods characterize the surface reasonably well. The spectrum exhibits a broad, flat peak over the range of 70 nm–3 μm . This is a range of two orders of magnitude and provides no distinct description of the surface. Both the ACVF and variational R_q identify a characteristic length, although the two lengths differ markedly. Variational R_q identifies a characteristic length of 50–80 nm, which is considerably larger than the apparent grain size of 20–30 nm. The length identified from half of the width of the peak in the ACVF is approximately 15–25 nm, which agrees very well with the apparent grain size. The ACVF, therefore, appears to identify characteristic lengths very well.

The advantage of using variational R_q is that roughness is directly evident at all lengths of the image. Roughness may be extracted from the PSD either by integrating the spectrum or by qualitatively comparing wavelength powers. Roughness may also be determined from the ACVF, but only for the length of the full image. At smaller length scales, roughness determination is problematic. An additional distinction between the treatments is the information density as a function of length. In the PSD, only ten data points describe the range of 0.1–1 L_i and the remaining 245 data points describe wavelengths smaller than 0.1 L_i . Since variational R_q is calculated in real space using linear steps, 90% of the data describes 90% of the image range.

V. CONCLUSIONS

With the proliferation of accessible and user-friendly SPMs, materials characterization with these tools is becoming routine. The ability to quantify surface topographic structure in a way that can be related to structure and processing is one of the great benefits of the techniques. Misrepresentation of structure can be avoided by using variational roughness calculations and explicitly considering criteria for valid quantification. For surfaces well represented by a Gaussian height distribution, variational rms is most applicable. When a surface is well represented by a superposition of sinusoidal relations, power spectral density is most appropriate. For many surfaces both/either will work. Both yield characteristic feature lengths; both can be used to extract fractal dimensions. Background subtractions, like any filtering routines, must be used in an objective and consistent manner. The plane associated with sample tilt must be determined from

images a minimum of ten times larger than the maximum lateral feature length or from an image size determined from composite variational analysis to be $>L^*$. The sample tilt plane rather than the image plane must be used on any smaller image or subset to avoid compromising the results. Least error in plane subtraction occurs in data at $L/L_i < 0.25$. 2D ACVF allows quantification of drift, tip radius, and tip related artifacts.

In summary, any roughness characterization is incomplete without a description of the measurement length, the size of the image in relation to the largest feature size, and the type of image processing performed on the image data.

¹See, for example, *Scanning Tunneling Microscopy and Spectroscopy*, edited by D. A. Bonnell (VCH, New York, 1993).

²M. W. Mitchell and D. A. Bonnell, *J. Mater. Res.* **5**, 2244 (1990).

³S. C. Langford, Ma Zhenyi, L. C. Jensen, and J. T. Dickinson, *J. Vac. Sci. Technol. A* **8**, 3470 (1990).

⁴D. M. Kulawansa, L. C. Jensen, S. C. Langford, and J. T. Dickinson, *J. Mater. Res.* **9**, 476 (1994).

⁵Y. L. Tsai and J. J. Mecholsky, Jr., *J. Mater. Res.* **6**, 1248 (1991).

⁶D. R. Denley, *Ultramicroscopy* **33**, 83 (1990).

⁷J. D. Kiely and D. A. Bonnell, *Mater. Res. Soc. Symp. Proc.* **318**, 401 (1994).

⁸S. Nakahara, S. C. Langford, and J. T. Dickinson, *J. Mater. Res.* **10**, 2033 (1995).

⁹E. A. Eklund, E. J. Snyder, and R. S. Williams, *Surf. Sci.* **285**, 157 (1993).

¹⁰R. M. Feenstra and G. S. Oehrlein, *Appl. Phys. Lett.* **47**, 97 (1985).

¹¹L. Spanos, Q. Liu, E. A. Irene, T. Zetter, B. Hornung, and J. J. Wortman, *J. Vac. Sci. Technol. A* **12**, 2653 (1994).

¹²R. C. Salvarezza, L. Vázquez, P. Herrasti, P. Ocón, J. M. Vara, and A. J. Arvia, *Europhys. Lett.* **20**, 727 (1992).

¹³J. M. Williams and T. P. Beebe, *J. Phys. Chem.* **97**, 6255 (1993).

¹⁴J. M. Gómez-Rodríguez, A. Asenjo, R. C. Salvarezza, and A. M. Baró, *Ultramicroscopy* **42**, 1321 (1992).

¹⁵J. D. Kiely and D. A. Bonnell, *Acta Metall. Mater.* (submitted).

¹⁶M. Niwa, H. Iwasaki, Y. Watanabe, I. Sumita, N. Akutsu, and Y. Akutsu, *Appl. Surf. Sci.* **60/61**, 39 (1992).

¹⁷Ph. Dumas, B. Bouffakhreddine, C. Arma, O. Vatel, E. Andre, R. Galindo, and F. Salvan, *Europhys. Lett.* **22**, 717 (1993).

¹⁸S. Talibuddin and J. P. Runt, *J. Appl. Phys.* **76**, 5070 (1994).

¹⁹J. S. Villarrubia, *Surf. Sci.* **321**, 287 (1994).

²⁰J. Tersoff and D. R. Hamann, *Phys. Rev. Lett.* **50**, 1998 (1983), *Phys. Rev. B* **31**, 805 (1985).

²¹G. Reiss, J. Vancea, H. Wittmann, J. Zweck, and H. Hoffmann, *J. Appl. Phys.* **67**, 1156 (1990).

²²X. B. Zhou and J. Th. M. De Hosson, *J. Mater. Res.* **10**, 1984 (1995).

²³Digital Instruments, *Nanoscope III Operator's Manual*.

²⁴V. Y. Milman, A. Stenashenko, and R. Blumenfeld, *Prog. Mater. Sci.* **38**, 425 (1994).

²⁵T. R. Thomas, *Rough Surfaces* (Longmans, London, 1982).

²⁶L. Spanos and E. A. Irene, *J. Vac. Sci. Technol. A* **12**, 2646 (1994).

Error budget of step height and pitch measurement using a scanning tunneling microscope with a three-dimensional interferometer

Toru Fujii,^{a)} Masataka Yamaguchi, and Masatoshi Suzuki
Nikon Corporation, 1-10-1, Asamizodai, Sagamihara, Kanagawa 228, Japan

Hirofumi Yamada
Department of Electronic Science and Engineering, Kyoto University, 606 Kyoto, Japan

Kan Nakayama
National Research Laboratory of Metrology, 1-1-4 Umezono, Tsukuba, 104 Ibaraki, Japan

(Received 8 September 1996; accepted 17 March 1997)

Scanning tunneling microscopy (STM) is well known for its high lateral resolution and nondestructive surface profiling capability but its potential as a profiler for dimensional measurement has rarely been exploited because of its unreliable vertical accuracy. In NANO III, a STM was presented that was capable of 250 μm line scanning secured by a parallel spring mechanism in the X direction; it employed an interferometric monitoring STM tip, allowing feedback controlled motion in the Z direction. For this device, errors of step height and pitch measurements on a surface topography standard are estimated. The uncertainty of the 3 μm pitch and 100 nm step height measurements using the STM is 1.05 and 0.41 nm, respectively. The uncertainties are induced by thermal expansion, the motion of the scanners, the nonlinearity of the interferometer, and geometrical error. In the 100 nm, step height measurement, the sum of the standard deviation of the measured data and the total error is still in the subnanometer range.

© 1997 American Vacuum Society. [S0734-211X(97)05504-2]

I. INTRODUCTION

The development of nondestructive surface profiling techniques with ultrahigh precision for large objects is finding increased interest for use in optical components and in semiconductor device manufacturing. Semiconductor devices microfabricated on silicon wafers, the standard size of which is rapidly shifting from 6 to 8 in., are also in need of such profiling techniques suitable for accurate linewidth and step height measurements.

Scanning tunneling microscopy (STM) and atomic force microscopy (AFM) are both new and emerging technologies and are being rapidly incorporated into profiling techniques for optical components and semiconductor devices, because these instruments have the capability of profiling microstructures with subnanometer lateral and vertical resolutions. Estimation of the errors of the measured results is essential, however, so that the potential of these devices can be exploited to the maximum for use in metrology.

Some groups have used interferometers to precisely monitor motions of the sample stage because the commonly used piezoelectric ceramic actuators have inherent problems of image distortion such as nonlinear response, hysteresis, creep, and interference among X , Y , and Z directional motion. Such distortion may be tolerable in applications for semiquantitative topographical studies but must be avoided in high precision measurements. Stemmer *et al.*¹ developed a STM with a tube scanner equipped with a two-dimensional (2D) interferometer for biological samples. Yamada *et al.*²

utilized a monolithic parallel spring mechanism as well as 2D interferometry to produce pure orthogonal X , Y , and Z motion from lead zirconium titanate (PZT) ceramic actuators. Fujii *et al.*³ measured the feedback motion of a STM as well as XY scanning motions with an interferometer in real time. Mc Waid and Schneir^{4,5} are developing a calibrated AFM that has a lateral measurement uncertainty of 3 nm and a vertical uncertainty of 1 nm. In other efforts, a STM is being developed to measure areas as large as 50 mm² with 0.1 nm resolution and 1 nm accuracy.⁶ Some groups applied other techniques for displacement monitoring such as a capacitance gauge,⁷ light intensity on a photodiode,⁸ and so on to correct the apparent distortion caused by PZT motion.

In this study, we measured the error motion of scanners, environment disturbances, and estimated uncertainties of pitch and step height values that do not include tip bending and contamination effects, using a STM with a three-dimensional interferometer.

II. INSTRUMENTATION

The STM used for this study, one equipped with a high precision three-dimensional interferometer, is described elsewhere.^{3,9} Figure 1(a) shows a schematic drawing of the STM. As shown in Fig. 1(b), a parallel spring scanner is used for one-dimensional 250 μm scanning. The tube scanner attached to it is equipped with a three-dimensional interferometer. To evaluate the accuracy of the STM measurement after interferometric calibration, we measured a surface topograph standard (VLSI Standards STM-1000A No. 2344-009-023), which has a 100 nm depth and a 3 μm pitch. Typical data measured with this STM are shown in Fig. 2. The parallel

^{a)}Present address: Swiss Federal Institute Lausanne, 1015 Lausanne, Switzerland; Electronic mail: toru.fujii@imt.dmt.epfl.ch

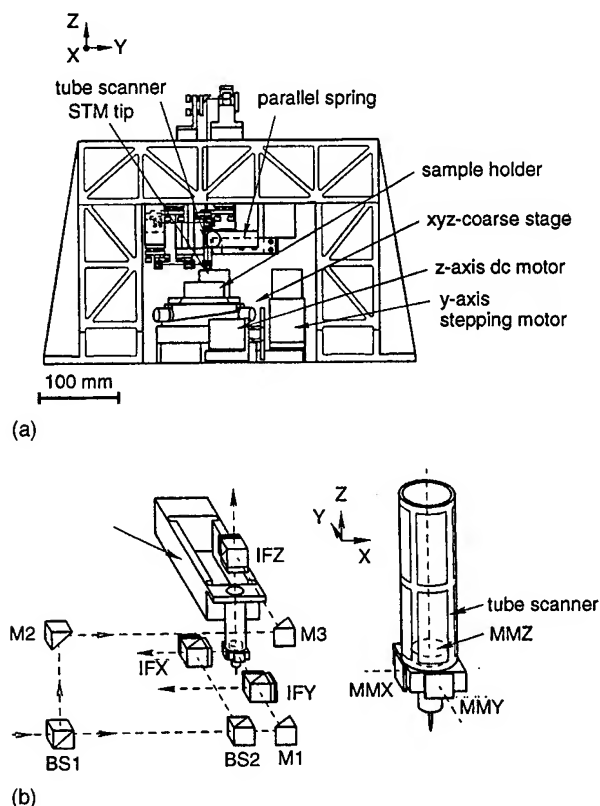


FIG. 1. (a) Schematic drawing of a STM. (b) Schematic drawing of a three-dimensional interferometer. BS1, 2—Nonpolarizing beam splitters; M1, 2, 3—steering mirrors; IFX, Y, Z—polarizing beam splitters with reference mirror and wave plates. The tube scanner has eight segmented outer electrodes to reduce the angular motion.

spring was used for line scanning in the X direction and the PZT tube scanner was used to drive the tip in the Z direction. A slow scan speed (300 nm/s) was used for our pitch and depth measurements because the feedback response, which is important in reducing overshoot effects, calculated from the lowest resonant frequency in the Z direction was as low as 760 Hz. An average of the measured step height with a standard deviation of the standard was 100.31 ± 0.44 nm, the total number of calculations was 36, and the pitch was 3.027 ± 0.016 μm when the total number was 33. On the other

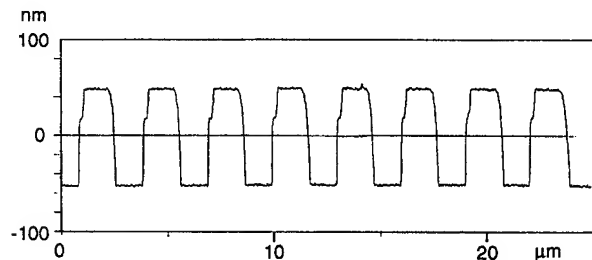


FIG. 2. Line scan data of a surface topograph standard. The parallel spring was used for line scanning and the tube scanner was used only for feedback motion in the z direction. The scale was calibrated by interferometry.

hand, the certificate values for the step height and the pitch were 97.8 ± 1.5 nm and 3.00 ± 0.02 μm , respectively.³

III. RESULTS

A. Causes of errors

1. Thermal drift

Using the three-dimensional interferometer for correcting the motion of the scanner, the relative positions of the reference mirror and the sample (and of the measurement mirror and probe) should be kept constant within the accuracy required. A major source of drift can be the slow temperature variations of a scanning probe microscope (SPM). For accurate measurement, reducing the cause of temperature variation is most important. Nanometer scale measurement using the STM with a three-dimensional interferometer typically requires temperature stability of better than 0.01 K/h. The thermal stability measured in the room used was 5 mK/h without any thermal feedback control.

Since drift between the reference mirror and the measurement mirror can be measured by the interferometer, thermal expansion in the vertical direction was calculated using the length and thermal expansion coefficients of each component of the instrument. The expansion coefficient obtained in the Z direction was 2.8×10^{-6} m/K. Thermal expansion of the pitch direction is also calculated in the same manner, and yielded a value in the Z direction of 0.11×10^{-6} m/K. The averaged slow scan speed was 300 nm/s, but we assumed that the lowest speed was 100 nm because the hysteresis changes the scan speed. We used a value of 30 s for single 3 μm pitch or step height measurements.

2. Motion error

Scanners do not produce a completely linear motion; they usually have a slightly angular motion as well, causing an Abbe error.¹⁰ The maximum angular motion of the parallel spring scanner measured by an autocollimator was 2.3 μrad over the range of 250 μm measured, i.e., better than 10 nrad/ μm . The Abbé offset between the laser beam line for the X displacement measurement and the apex of the tunneling tip was 15 mm, resulting in an Abbé error of 35 nm per 250 μm . For the Z direction, the maximum offset was 1 mm and the error was 2.3 nm per 250 μm .

The angular motion of the tube scanner was measured using the optical lever method. The reflected beam position changes according to the angular motion of the measurement mirror. The beam position was measured by a four-segmented photodiode (spot 4D). The angular motion of the tube scanner when actuating in the Z direction was 1.4×10^{-5} rad over 8000 nm and 1.7×10^{-7} rad/100 nm. The tilting motion of the mirror in the Z direction during its 3 μm travel in the X direction was 30 nm.

3. Geometric error

The combined machining and assembling error is typically better than 0.1° (1/600 rad), and the sample setting error, which occurs only in the horizontal direction, is within 1° (1/60 rad).

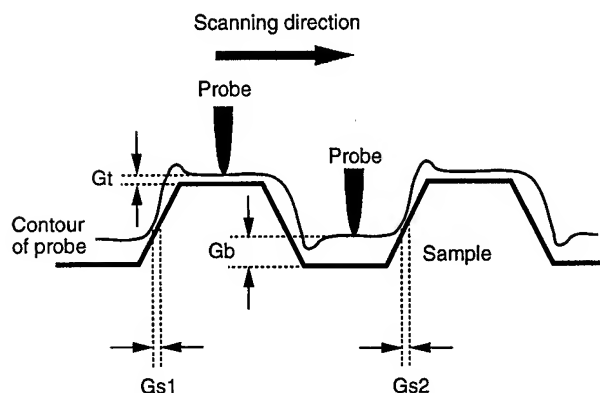


FIG. 3. Exaggerated error caused by the feedback system and the tip shape. G_t —Vertical interaction gap at the top of the pattern; G_b —vertical interaction gap at the bottom; G_{s1} , 2—lateral interaction gaps at the side wall.

4. Feedback error

The interaction gap fluctuates especially around a step edge because the feedback response is finite. This fluctuation has to be estimated when a sample has a complicated shape. A contour of the probe scanned on a periodic pattern is shown in Fig. 3. The overshoot, undershoot, and deviation of the contour are exaggerated. G_{s1} and G_{s2} are lateral gaps at half-maximum. G_t and G_b are gaps at the top and the bottom, respectively. The pitch measurement of the cyclic pattern is not affected by the feedback response because the repeated feedback motion can maintain the same gap at every half-maximum point ($G_{s1} = G_{s2}$). The gap distance depends on the change of probe height in the linear feedback control system. In this experiment, the unity gain of the feedback system was given at about 300 Hz. The effective gain of this measurement was 10^3 at 0.3 Hz, and deviation of the tunneling gap between the top and bottom of the pattern (i.e., $G_b - G_t$) was $100 \text{ nm}/10^3 = 0.1 \text{ nm}$.

5. Interferometer error

It is well known that interferometer measurements utilizing polarization show a small periodic error of the phase; here, 20 nmp-p was measured without using a balanced detection method.⁸ This periodic error was reduced to 1.00 nmp-p, i.e., 0.35 nm rms, by balanced detection as calculated in Ref. 9.

6. Tip shape error

A pitch and step height measurement is the most suitable for estimating the performance of a SPM for a dimensional measurement because the measurement of cyclic patterns seen in Fig. 3 is free of the tip shape problem. This problem has been discussed since Stedmann.¹¹

B. Height and pitch measurements

The standard deviation of the measured step height discussed in Sec. II suggests that even if there were a contamination layer thickness, it would be uniform. Contamination

TABLE I. Errors in the step height measurement (in nm).

Abbé error of the parallel spring	0.03
Error of the tube scanner	0.18
Geometric error	0.014
Thermal drift	0.12
Interferometer error	0.35
Total (rms)	0.41
(Std. deviation of measurement)	(0.44)

does not change the value of step height and pitch. Moreover, a STM tip modulation of 0.35 nm in the Z direction could change a tunneling current of one decade in this experiment. The displacement was very accurate because it was measured by the interferometer, and suggested that condition of the tunneling junction was good. We assume that contamination effects⁵ are negligible in our measurements.

The shape of the mechanical ground STM tip used in this experiment was observed by scanning electron microscopy (SEM). The SEM photograph showed that the tip stiffness was more than 100 times harder than the contamination stylus for AFM,¹² which bent less than 5 nm. We also assume that the tip bending effect^{12,13} was negligible.

Tables I and II show the total errors calculated in the height and the pitch measurements, respectively. Two of the total errors were calculated based on the sum of the square of the individual errors by means of the root mean square value method. In the pitch measurement, the standard deviation of the measured data is dominant; however, in the step height measurement, both are comparable and a sum of the two is 0.60 nm rms.

C. Comparison with other methods

To evaluate differences between the STM data and the certificate value of the standard pattern, step height and pitch of the standard were measured using a SEM (Akashi Beam Technology Co. CL-130) and a Mirau interferometer.¹⁴ Values of the pitch measured by the SEM and the Mirau interferometer were 3.12 ± 0.04 and $3.08 \mu\text{m}$, respectively. The resolution of the SEM photograph was about 10 nm. Measurement between half-maximum points, which is the same method used in the STM measurement, was used for the Wyko, image. Because the sampling spacing of the Mirau interferometer is as large as $0.2 \mu\text{m}$, the measured pitch was 1/10 for 10 line-and-spaces. The distance of the 10 line-and-spaces was measured four times and was $30.8 \mu\text{m}$ each time.

The depth was also measured from the Mirau interferometer image. A value of $100.97 \pm 1.45 \text{ nm}$ was obtained from

TABLE II. Errors in the pitch measurement (in nm).

Abbé error of the parallel spring	0.45
Error of the tube scanner	0.77
Geometric error	0.42
Thermal drift	0.005
Interferometer error	0.35
Total (rms)	1.05
(Std. deviation of measurement)	(16)

19 data runs. The digit resolution in this direction of the Mirau interferometer was 0.01 nm. The flat part of the bottom of the pattern could not be observed in the Wyko image whereas it could be clearly seen in the STM image. A comparison of the methods for pitch and step height follows.

Pitch

STM:	$3.027 \pm 0.016 \mu\text{m}$, $n = 33$, and uncertainty of 1.05 nm;
Wyko:	$3.08 \pm 0.00 \mu\text{m}$, $n = 4$ (4×10 pitches);
SEM:	$3.12 \pm 0.04 \mu\text{m}$, $n = 25$;
certificate value:	$3.00 \pm 0.02 \mu\text{m}$.

Step height

STM:	$100.31 \pm 0.44 \text{ nm}$, $n = 36$, and uncertainty of 0.41 nm;
Wyko:	$100.97 \pm 1.45 \text{ nm}$, $n = 19$;
certificate value:	$97.8 \pm 1.5 \text{ nm}$.

The SEM and Wyko measurements indicated that the values measured by both instruments were larger than the certificate values.

IV. CONCLUSION

Errors of the step height and pitch measurements with the STM equipped with a high accuracy interferometer were estimated. The uncertainty of the 3 μm pitch and 100 nm step

height measurements was 1.05 and 0.41 nm, respectively. The uncertainties are caused by thermal expansion, motion of the scanners, the nonlinearity of the interferometer, and geometrical error. In the step height measurement of 100 nm, the sum of the standard deviation of the measured data and the total error is still in the subnanometer range.

¹A. Stemmer, A. Engel, R. Häring, R. Reichelt, and U. Aebi, *Ultramicroscopy* **25**, 171 (1988).

²H. Yamada, T. Fujii, and K. Nakayama, *Jpn. J. Appl. Phys.* **1** **28**, 2402 (1989).

³T. Fujii, M. Yamaguchi, and M. Suzuki, *Rev. Sci. Instrum.* **66**, 2504 (1995).

⁴T. H. McWaid and J. Schneir, Technical Program of Industrial Applications of SPM, Workshop of NIST, March 1994, p. 81; J. Schneir, T. H. McWaid, J. Alexander, and B. P. Wilfley, *J. Vac. Sci. Technol. B* **12**, 3561 (1994).

⁵J. Schneir, J. S. Villarrubia, T. H. McWaid, V. W. Tsai, and R. Dixon, *J. Vac. Sci. Technol. B* **14**, 1540 (1996).

⁶E. C. Teague, *J. Vac. Sci. Technol. B* **7**, 1898 (1989).

⁷L. Libioulle, A. Ronda, M. Taborrelli, and J. M. Gilles, *J. Vac. Sci. Technol. B* **9**, 655 (1991).

⁸R. C. Barrett and C. F. Quate, *Rev. Sci. Instrum.* **62**, 1393 (1991).

⁹T. Fujii, M. Suzuki, M. Yamaguchi, R. Kawaguchi, H. Yamada, and K. Nakayama, *Nanotechnology* **6**, 121 (1995).

¹⁰E. C. Teague and C. Evans, *Tutorial Notes* (ASPE, Raleigh, NC, 1988), p. 99.

¹¹M. Stedman, *J. Microsc.* **152**, 611 (1988).

¹²T. Fujii, M. Suzuki, M. Miyashita, M. Yamaguchi, T. Onuki, H. Nakamura, T. Matsubara, H. Yamada, and K. Nakayama, *J. Vac. Sci. Technol. B* **9**, 666 (1991).

¹³P. Markiewicz and M. C. Goh, *Rev. Sci. Instrum.* **66**, 3186 (1995).

¹⁴Wyko TOPO3D, 80 μm cut off filtered, 40 \times objective, measured cross section length of about 250 μm .

Principle of atomic grating and its application in nanotechnology

H. W. Hao, H. Qiao, H. K. Cheng, W. W. Wei, and W. K. Yi

Department of Precision Engineering, University of Science and Technology of China, 230026 Hefei, China

(Received 12 September 1996; accepted 31 March 1997)

We present here a new method using atomic grating fringe to achieve resolution as high as one-tenth of an atomic lattice in nanometrology system. The principle of this method is as follows: by superposing two scanning probe microscope (SPM) atomic images, when one is rotated to a specific small angle, moire fringes will be observed. These moire fringes have symmetry and larger periodicity than atomic lattice. Comparing the difference of the fringes obtained before and after sample displacement, we can determine the direction and quantity of movement of the sample within one-tenth of the lattice spacing. This method is suitable to characterizing the stability of the SPM instrument and to single atom location and manufacturing. © 1997 American Vacuum Society. [S0734-211X(97)04804-X]

I. INTRODUCTION

Since the first report of the scanning tunneling microscopy (STM) by Binnig *et al.*¹ the surprising ability of STM to reveal the surface topography in atomic scale has attracted much attention. Within the last few years, the techniques of the STM have made rapid progress, and they are applied to various fields, one of them is to nanometrology. STM was used to measure the surface roughness² and microhardness,³ the STM was used as a high resolution locating, and the atomic lattice obtained by STM was used as a scale to calibrate the other images.⁴ All these applications show the great potential in the field of nanometrology. But the application of STM to metrology is still in its infancy. For instance, in order to get accuracy on nanometer or atomic scale, the sensitivity of the measuring system must reach one-tenth of the accuracy. This problem is not perfectly solved until now.

In this article, we propose a subatomic measuring system using atomic grating. Atomic grating system consists of an STM and image processing techniques. We found that when two identical STM images of highly oriented pyrolytic graphite (HOPG) are rotated from each other for a specific small angle and then superposed, moire fringes will be observed over the usual atomic pattern of graphite. These moire fringes have hexagonal symmetry and larger periodicity than atomic lattice. When one above two images is moved rela-

tively to the other for one atomic spacing, moire fringes will move for one period along the direction perpendicular to the direction which image moves along. Moire fringes have function of magnifying the displacement. The magnification is dependent on the angle rotated. Based on this characteristic, atomic grating can reach the sensitivity of one-tenth of atomic resolution. This atomic grating system is suitable to characterizing the stability of the scanning probe microscopy (SPM) instrument and to single atom locating and manufacturing.

II. THEORY

When two plane waves, whose directions differ by a small angle of δ , are superposed, then a moire pattern can be observed. The superposition of these two plane waves can be written in the form:⁵

$$\begin{aligned} z &= \frac{1}{2} \cos\{2\pi(x \cos \frac{1}{2}\delta + y \sin \frac{1}{2}\delta)\} \\ &+ \frac{1}{2} \cos\{2\pi(x \cos \frac{1}{2}\delta - y \sin \frac{1}{2}\delta)\} \\ &= \cos\{2\pi x \cos \frac{1}{2}\delta\} \cos\{2\pi y \sin \frac{1}{2}\delta\} \\ &\approx \cos(2\pi x) \cos(2\pi y \frac{1}{2}\delta) \quad \text{if } \delta \ll 1. \end{aligned} \quad (1)$$

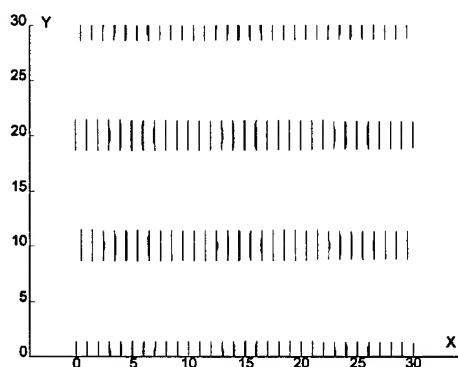


FIG. 1. The contours of two plane waves of Eq. (1), where $z=0.9$ and $\delta=0.1$.

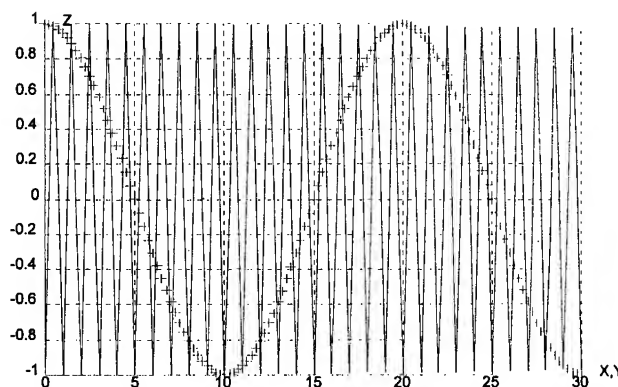


FIG. 2. Cross sections of Fig. 1 at $X=10$ marked with plus signs, and at $Y=10$ in solid lines.

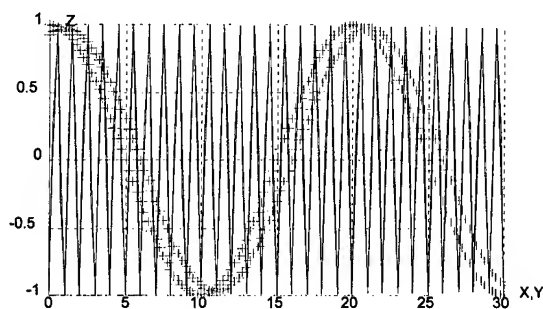


FIG. 3. Profiles of Eq. (2) at $X=10$ marked with plus signs, and $Y=10$ in solid lines where $\delta=0.1$. The displacement Δx equals 0 and 0.1, respectively.

Here the wave number vectors of the two form angle $+\delta/2$ and $-\delta/2$ with the x axis, respectively. Figure 1 shows the 0.9 contours of Eq. (1), where $\delta=0.1$. Figure 2 is the cross section of Fig. 1 at $x=10$ and $y=10$. From Figs. 1 and 2, we can see when two waves with directions which differ by 0.1 rad superposed, moire fringes will be produced. The period of moire patterns is $\lambda=2/\delta$ and the axis of the moire pattern is perpendicular to wave number vectors of the plane waves. But if we consider the average effect in a small window, the period of moire patterns will be $\lambda=1/\delta$.

From Eq. (1), it can be seen that along the y axis the moire pattern $\cos(2\pi y/2\delta)$ appears. The period of the moire pattern is $2/\delta$ while the plane waves' period is 1. In Eq. (1), if the second plane wave has a displacement, say Δx , then Eq. (1) should be written as

$$\begin{aligned} z &= \frac{1}{2} \cos\{2\pi(x \cos \frac{1}{2}\delta + y \sin \frac{1}{2}\delta)\} \\ &\quad + \frac{1}{2} \cos\{2\pi[(x + \Delta x) \cos \frac{1}{2}\delta - y \sin \frac{1}{2}\delta]\} \\ &= \cos\{2\pi x \cos \frac{1}{2}\delta + \pi\Delta x \cos \frac{1}{2}\delta\} \\ &\quad * \cos\{2\pi y \sin \frac{1}{2}\delta - \pi\Delta x \cos \frac{1}{2}\delta\} \\ &\approx \cos(2\pi x + \pi\Delta x) \cos(\pi\delta y - \pi\Delta x) \quad \text{if } \delta \ll 1. \end{aligned} \quad (2)$$

From Eq. (2), it is found that when one plane wave has a displacement of Δx , the moire pattern has a displacement of $2\Delta x/\delta$. Figure 3 shows when one plane has a small displacement of 0.1, and two planes have a slight angle 0.1, the moire pattern displacement is 2, which is more than that of the plane wave. It suggests that moire fringes have the function of magnifying displacement. Also if we consider the

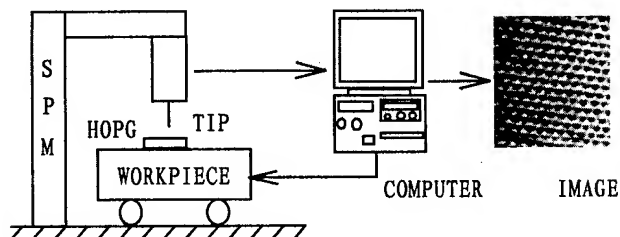


FIG. 4. Schematic of the atomic grating system based on STM.

Combined STM Pattern, no displacement

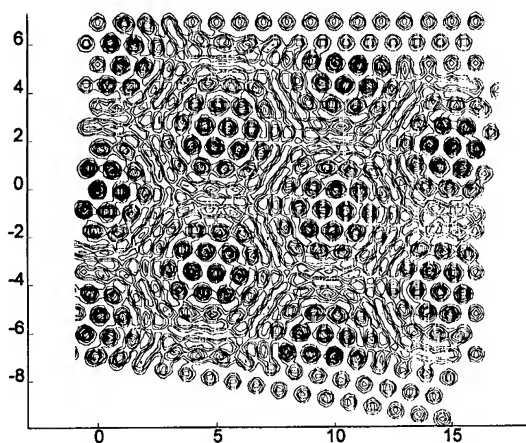


FIG. 5. Simulation result of superposing of two HOPG images. The reference image has rotated 9.8° and there is no displacement of the second image. The numbers on axes represent atoms.

average effect in a small window, the moire pattern displacement will be $\Delta x/\delta$ when one plane wave has a displacement of Δx .

The atomic grating system consists of an STM and makes use of image processing techniques. Figure 4 is the schematic of the atomic grating system. In engineering practice, only one piece of HOPG is needed to be attached to a movable workpiece. Before the workpiece moves, a STM image is obtained first and then rotated for a specific angle. This rotated image will be stored as the reference. While the workpiece moves, another new STM image is obtained and added to the reference STM image. By calculating the change of distance of the two moire fringes, we can know the quantity of the workpiece's movement. Because moire fringes have the function of magnifying displacement, for instance by the factor 10, we can achieve subatomic resolution measurement by atomic grating.

III. COMPUTER SIMULATION

A HOPG liked image was generated by computer, and then rotated to a specific angle. This image was used as a

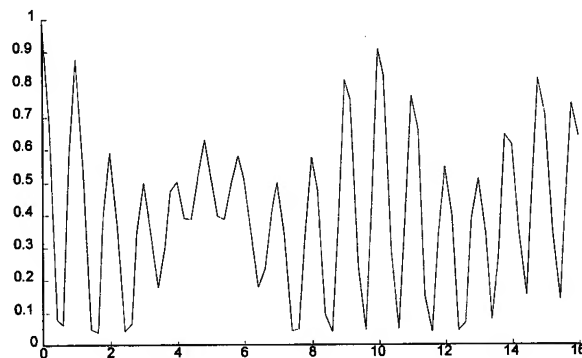


FIG. 6. Cross section of Fig. 5 at $Y=0$. The numbers on the horizontal axis represents atoms, and the numbers on the vertical axis are arbitrary units.

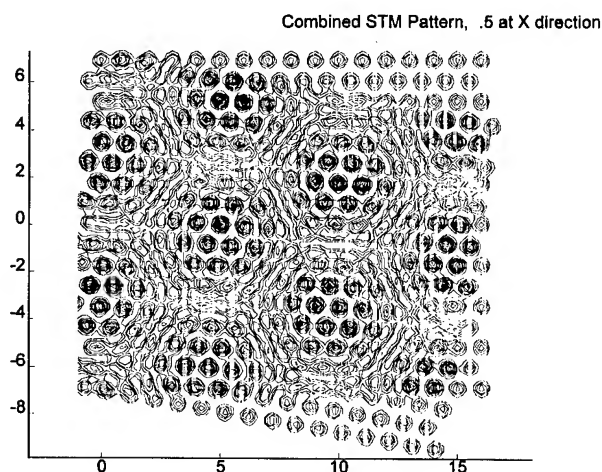


FIG. 7. Simulation result of superposing of two HOPG images. The reference image has rotated 9.8° and the second image has a displacement of 0.5 atomic lattice spacing in the x direction. The moiré fringe has a displacement of three atomic lattice spacings in the y direction. The numbers on axes represent atoms.

reference image. Then the second HOPG like image was generated and superposed to the reference image. Moiré fringes would be observed over the usual atomic pattern of graphite, as shown in Fig. 5. The numbers on axes represent atoms. This moiré fringes have hexagonal symmetry and larger periodicity than atomic lattice. The profile was shown in Fig. 6. This moiré fringes indicate the initial position of the workpiece. When the second HOPG image was moved half lattice in x direction and then superposed to the reference image, the second moiré fringes would be gotten, as shown in Fig. 7. The whole pattern moves in y direction by half of big period. Its profile along $Y=0$ was shown in Fig. 8. Comparing these two moiré fringes' changes, the movement of the workpiece can be obtained in subatomic sensitivity.

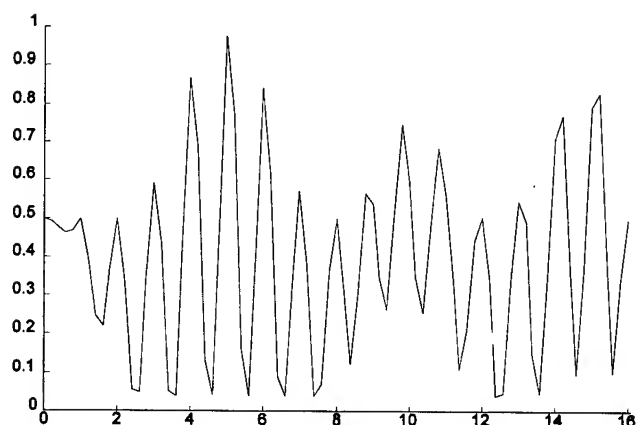


FIG. 8. Cross section of Fig. 7 at $Y=0$. The numbers on the horizontal axis represents atoms, and the numbers on the vertical axis are arbitrary units.

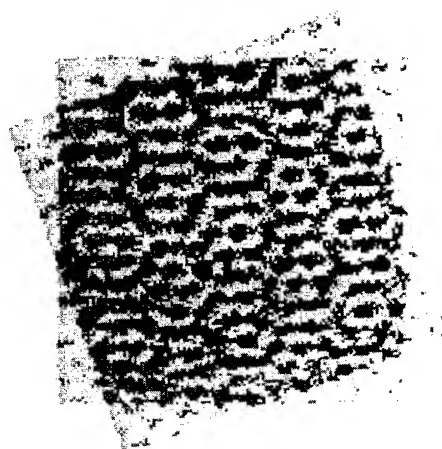


FIG. 9. Moiré fringes of superposing of two real STM images in the initial position, where the relative rotation angle is 19.6° .

IV. EXPERIMENT

Our experiment was done with a homemade STM. When two STM images of HOPG were rotated clockwise and anticlockwise relative to axis X of the angle of 9.8° , respectively, we got images A and B. Images A and B were superposed. Moiré fringes resulted, see Fig. 9. Then image A was moved a displacement of four pixels along x direction and superposed to image B. Here the length of 50 pixels equals 1 nm. The second moiré fringes were got in Fig. 10. We used an image processing method, which we call the template-matching method (TMM), to calculate the displacement of the moiré fringes between Figs. 9 and 10. We first selected a window W_a of a specific size in Fig. 9 as a reference window, and another window W_b of the same size in Fig. 10. The similarity between W_a and W_b was defined as C_{mn} in Eq. (3):

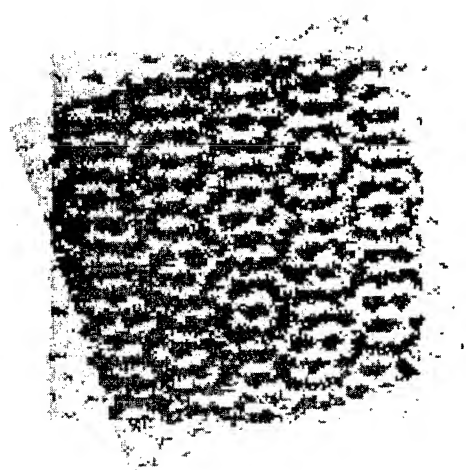


FIG. 10. Moiré fringes of superposing two real STM images at the second position, where one image has a displacement of four pixels along horizontal direction. The moiré fringes have a movement of 12 pixels along the vertical direction.

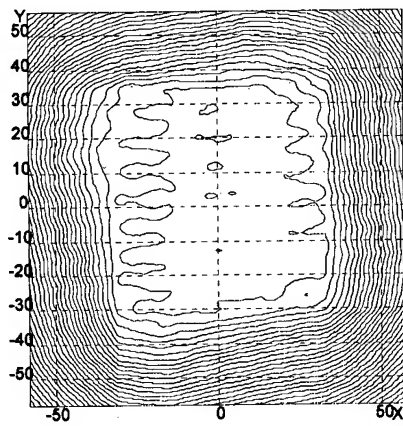


FIG. 11. C_{mn}^{rs} contour of the TMM method, where the units of the x and y axis are both pixels and the length of 50 pixels is 1 nm. The global minimum locates at (0,12).

$$C_{mn} = \sqrt{\frac{1}{n} \sum_{i,j} (Z_{aij} - Z_{bij})^2}, \quad (3)$$

where Z_{aij} and Z_{bij} are pixel values in Wa and Wb, respectively, m and n are coordinates of Wbs center. C_{mn} 's contour was shown in Fig. 11, where the units of the x and y axis are both pixels. From Fig. 11, we could see that C_{mn} 's global minimum located at (0,12). It implied that the displacement

of the moire patterns moved 12 pixels along the y direction. Thus the magnification of the moire fringes was three. This agreed with Eq. (2).

V. DISCUSSION

In this article, we proposed the idea of atomic grating. TMM method was used to identify the quantity and direction of images' shift at subatomic sensitivity. It can be used in characterizing the stability of SPM instrument and single atom location system. With a feedback control system, the thermal shift can be compensated. The experiment result was in agreement with the theory. Of course, several key techniques should be well solved before atomic grating's being put into engineering application. There are lots of sources which affect SPM measurement, such as noise, thermal drift, and piezo scanner imperfections. Usually piezo scanners are neither linear nor stable, but we can use available calibration and compensation techniques to minimize errors below 2%. As well, grating averaging effect makes the measurement not very sensitive to the errors mentioned above.

¹G. Binning, H. Rohrer, Ch. Gerber, and E. Weibel, Phys. Rev. Lett. **49**, 57 (1982).

²N. Garcia, A. M. Baro, and R. Miranda, Metrologia **21**, 135 (1985).

³K. Kanazawa et al., J. Jpn. Soc. Heat Treat **29**, 167 (1989).

⁴H. Kawatatsu et al., J. Vac. Sci. Technol. B **12**, 1681 (1994).

⁵T. Mori, H. Abe, K. Saiki, and A. Koma, Jpn. J. Appl. Phys. **32**, 2945 (1993).

Size and arrangement of elementary fibrils in crystalline cellulose studied with scanning tunneling microscopy

Y. Z. Zhang

Institute of Microbiology, Shandong University, Jinan 250100, China

X. L. Chen

Cotton Research Centre, Shandong Academy of Agricultural Science, Jinan 250100, China

J. Liu and P. J. Gao

Institute of Microbiology, Shandong University, Jinan 250100, China

D. X. Shi and S. J. Pang^{a)}

Laboratory of Vacuum Physics, Center for Condensed Matter Physics, Chinese Academy of Sciences, P.O. Box 2724, Beijing 100080, China

(Received 8 September 1996; accepted 12 May 1997)

Scanning tunneling microscopy (STM) was used to investigate the ultrastructure of cellulose. The materials used in the experiments were cotton fiber, dewaxed cotton fiber, and microcrystalline cellulose. The results showed that the elementary fibrils in all these kinds of cellulose could be directly observed from the surface and cross-sectional view with high resolution. The elementary fibrils assembled together in parallel, and their lateral dimension showed a great variability in different kinds of cellulose, but was uniform in the same kind of cellulose. Elementary fibrils were the smallest structural units of cellulose, and they further aggregated into microfibrils, and the microfibrils constituted fibrils. In each gradation, the fibers piled up in parallel. STM was useful in studying the fine structure of cellulose. © 1997 American Vacuum Society. [S0734-211X(97)12704-4]

I. INTRODUCTION

Cellulose is the major polysaccharide component of plant cell wall and is the most abundant organic compound on the planet. It plays a very important role in the carbon and energy cycles of the biosphere. Thus, more and more attention has been paid to studying its structure, function, and biodegradation.¹ The existence of an elementary cellulose fibril (ECF), as the structural units of cellulose, was proposed a long time ago.^{2,3} Since then, the size of the smallest crystalline unit has been measured in cellulose from valonia,⁴⁻⁸ bacteria,⁴⁻⁸ cotton,^{2,7,9} flax,² jute,² and ramie,^{2,4-8,10,11} by transmission electron microscopy (TEM), wide-angle x-ray scattering (WAXS), and small-angle x-ray scattering (SAXS). It is evident from the reported data that the size of these crystalline regions showed a great variability. In some cases, this was interpreted as being due to the agglomeration of elementary subunits of about 35 Å in size,¹¹ and in other cases this interpretation was refuted and the existence of an ECF was questioned altogether.⁸ However, Jakob *et al.* reported that the wood cell wall was built with elementary cellulose fibrils having a uniform thickness of 25 ± 2 Å, which was shown by investigating the same sample independently with three different experimental techniques, TEM, WAXS, and SAXS.⁹ But to date, to the best of our knowledge, the ultrastructure of cellulose has not yet been known completely.

The appearance of the scanning probe microscopy (SPM) technique has principally opened new possibilities for the

studying of biological objects, such as macromolecules. Compared with other structure analytical tools, SPM can be performed in ambient condition, and its resolution is very high. It has been widely used to image the topography of DNA, protein, and membranes.¹² In the past, atomic force microscopy has been used to study the structure of algal cellulose.^{13,14} Here, scanning tunneling microscopy (STM) was used to investigate the ultrastructure of three kinds of cellulose, and high-resolution STM images have been obtained.

II. MATERIALS AND METHODS

Cotton fiber, dewaxed cotton fiber, and microcrystalline cellulose were used as the materials. The cotton fiber was from the cultivated cotton variety H123 A. The dewaxed fibers and microcrystalline cellulose Sigma cell type 50 were purchased as commercial reagents. The cotton fiber and dewaxed cotton fiber were made into powder by smashing, washed with distilled water, and then heated to dry at 50 °C. For the STM experiments, the cellulose powder was suspended in anhydrous ethanol, 20 µl of suspension was dropped on newly cleaved highly oriented pyrolytic graphite, and then dried in air. The STM experiments were carried out in ambient environment with a domestic STM setup CSPM-930a (manufactured by the Institute of Chemistry, Academia Sinica). The normal STM constant current mode was used, with tungsten tips made by electrochemical etching. Tunneling current and bias voltage were indicated in the relevant legends of the photographs.

^{a)}Author to whom correspondence should be addressed.

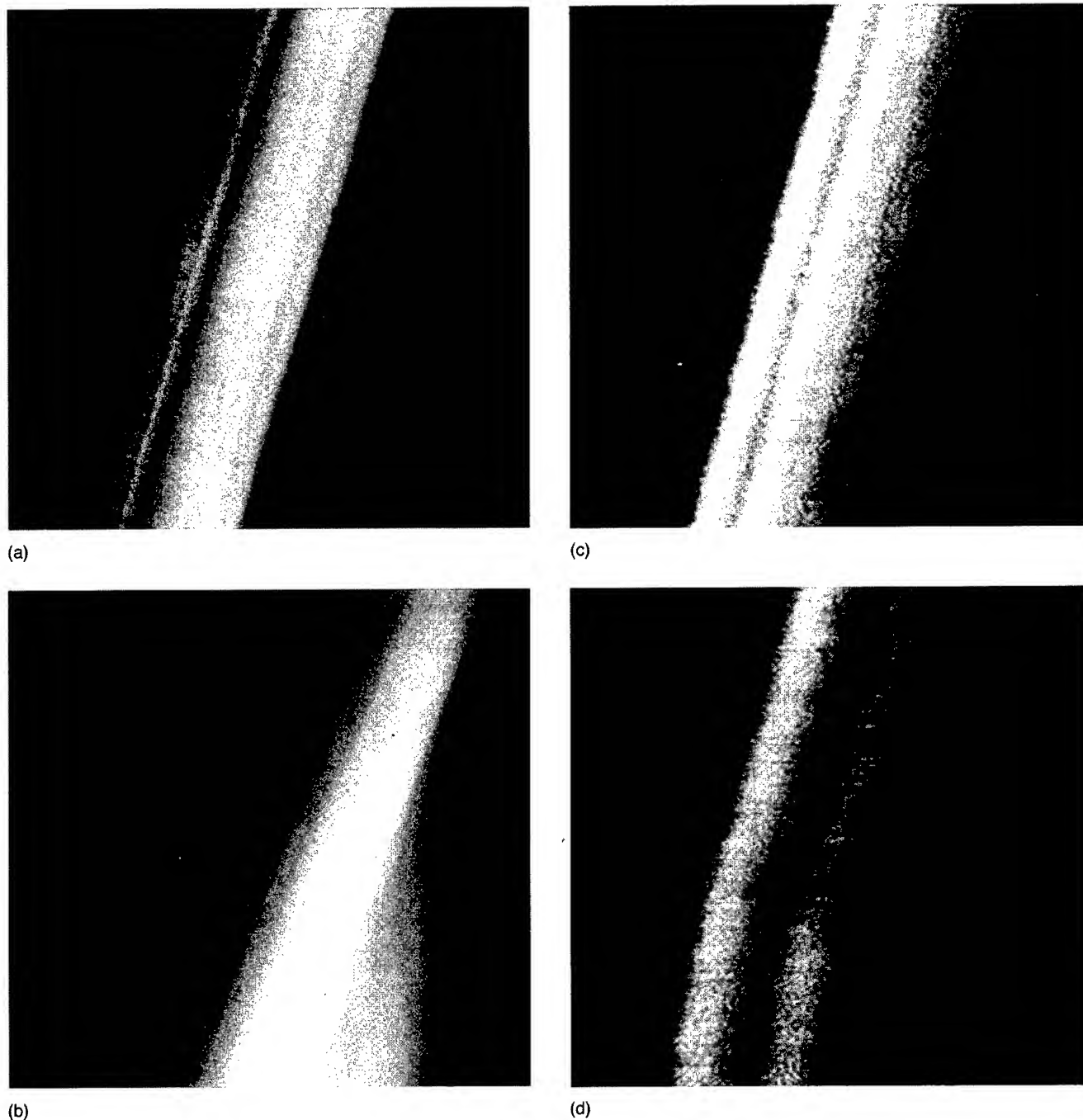


FIG. 1. The STM images of cotton fiber of the variety H123 A. $V_{\text{bias}} = 740$ mV, $I_{\text{ref}} = 0.18$ nA, scan area: (a) $320 \text{ nm} \times 320 \text{ nm}$, (b) $64 \text{ nm} \times 64 \text{ nm}$, (c) $32 \text{ nm} \times 32 \text{ nm}$, and (d) $16 \text{ nm} \times 16 \text{ nm}$.

III. RESULTS AND DISCUSSION

Cotton fiber is the only native pure cellulose material with an amount of cellulose up to 95%–97%. After being dewaxed, it is a good kind of substrate for structure and biodegradation researchers.

Because the dimension of the cotton fiber with 10–100 μm in diameter was greatly beyond the STM maximum scanning scale, the STM tip could only scan in a part area of the cotton fiber surface. Figure 1 shows the STM images of the H123 A cotton fiber. The cellulose fibrils were clearly

observed [Fig. 1(a)]. The lateral diameter ranged from 90 to 120 nm, and their length was greatly beyond the scanning scale. It could be also seen that the cellulose fibrils were composed of thinner fibers, which was called microfibril, with 25–35 nm in diameter. With higher resolution, the microfibrils could be seen more clearly [Fig. 1(b)]. It was also shown that the microfibrils consisted of much thinner fibers with a $25 \pm 3 \text{ \AA}$ diam. The thinner fibers were elementary fibrils, which were made up of cellulose molecules. Their structures were shown in Figs. 1(c) and 1(d). Some of the



FIG. 2. The STM image of the dewaxed cotton fiber, $V_{\text{bias}}=830$ mV, $I_{\text{ref}}=0.18$ nA, and scan area, $32\text{ nm}\times 32\text{ nm}$.



FIG. 3. The STM image of microcrystalline cellulose Sigma cell type 50. $V_{\text{bias}}=475$ mV, $I_{\text{ref}}=0.81$ nA. Scan area $40\text{ nm}\times 40\text{ nm}$.

elementary fibrils were agglomerated tightly, some of them were dissociated from others. It could be concluded that the cotton fiber was built with elementary fibrils. Cellulose, as the native macromolecule, was constituted in the stepping mode. The cellulose molecule was β -1, four-linked *D*-glucose monomers with cellobiose as repeat units; they aggregated together to form elementary fibrils depending on intra- and intermolecular hydrogen bonds. Elementary fibrils were the smallest structural units of cellulose. Elementary fibrils further assembled into microfibrils, and microfibrils constituted fibrils. In each gradation, the fibers were parallel to each other.

Figure 2 shows the STM image of dewaxed cotton fiber with high resolution. The surface structure of the dewaxed cotton fiber could be observed, and was similar to that of the H123 A cotton fiber; the elementary fibrils could be seen directly. The thinnest elementary fibril assembled in parallel with a diameter of 30 ± 3 Å, the chains in the elementary fibrils with a 7–10 Å diam might be the cellulose molecule.

In order to understand the size and arrangement of the elementary fibrils in the crystalline cellulose comprehensively, the surface and transverse section structures of the highly crystallized microcrystalline cellulose were studied with STM. The STM image is shown in the Fig. 3. It could be seen from the cross-sectional view that the elementary fibrils piled up, which was consistent with that observed from the surface view. The cross section of the elementary fibril was oval in shape, the diameters of the elementary fibrils were 18 ± 2 Å.

The elementary fibril diameters of these three kinds of cellulose used above are summarized in Table I. The results suggested that the diameter of the elementary fibrils in the same kinds of cellulose was uniform. This was consistent

with the results reported by Jakob *et al.* Otherwise, its actual value varied with the materials. It is evident from the reported data that the size of these crystalline regions showed a great variability. This might be due to the agglomeration of elementary subunits,¹¹ just like the observed results in this article.

The ultrastructure of cellulose has been studied with electron microscopy, and a lot of information about this complex biological system was gained. However, to date, to the best of our knowledge relatively little is known about its detailed structure at the microscopic level. This is, in part, due to the fiber structure of the cellulose with crystalline and amorphous domains at the microscopic level. Transmission electron microscopy requires tedious sample preparation procedures and staining methods, which contain the danger of altering the structure of the sample, it is difficult to directly observe the structure of the elementary fibrils with higher resolution. Small-angle x-ray scattering has widely been applied to the study of cellulose fibril, and some parameters, such as the diameter of elementary fibril, might be determined with higher accuracy.¹⁵ However, SAXS does not provide direct images of the structure, and the information gained with SAXS is averages. Compared with TEM and

TABLE I. The elementary fibril diameter measured in three types of celluloses.

Types of cellulose	Diameter of the elementary fibrils (Å)
Cotton fiber of H123 A	25 ± 3 Å
Dewaxed cotton fiber	30 ± 3 Å
Microcrystalline cellulose Sigma cell type 50	18 ± 2 Å

SAXS, the sample preparation for STM is very simple, and the STM now available can process data almost immediately using relatively compact equipment and can provide three-dimensional structural information with atomic resolution, so STM could be used to investigate the surface structure from real time and real space. From the results above, we can conclude that it is possible to directly observe the fine structure of cellulose of a crystalline region, such as the arrangement of microfibrils and elementary fibrils; it is especially possible to observe the structure and size of the individual elementary fibril and compare the differences between them. Moreover, STM is able to directly investigate the structure of biomolecules under physiological conditions; it can be applied to observing the dynamic process of biodegradation of cellulose by cellulase, such as the adsorption of cellulase on a cellulose surface.

IV. CONCLUSION

To sum up, STM was used to investigate the ultrastructure of cellulose of cotton fiber, dewaxed cotton fiber, and microcrystalline cellulose. The existence of an elementary fibril could be demonstrated by STM with high resolution. Moreover, the size and arrangement of the elementary fibrils in the cellulose could also be clearly observed. All the results showed that STM is a useful technique for investigating the fine structure of cellulose.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation of China (NSFC). The authors also give their warm thanks to Mr. Jun Zhang for providing the cotton fiber from the cultivated variety H123 A.

¹P. Beguin and J. P. Aubert, *FEMS Microbiol. Rev.* **13**, 25 (1994).

²A. N. Heyn, *J. Appl. Phys.* **5**, 519 (1955).

³A. N. Heyn, *J. Appl. Phys.* **9**, 1113 (1955).

⁴A. M. Scallan, *Text. Res. J.* **41**, 647 (1971).

⁵J. Haase, R. Hosemann, and B. Renwanz, *Kolloid Z. Z. Polym.* **251**, 871 (1973).

⁶J. Haase, R. Hosemann, and B. Renwanz, *Colloid Polym. Sci.* **252**, 712 (1974).

⁷E. K. Boylston and J. J. Hebert, *J. Appl. Polym. Sci.* **25**, 2105 (1980).

⁸H.-P. Fink, D. Hofmann, and H. J. Purz, *Acta Polym.* **41**, 131 (1990).

⁹H. F. Jakob, D. Fengel, S. E. Tschegg, and P. Fratzl, *Macromolecules* **28**, 8782 (1995).

¹⁰A. M. Hineleh and D. J. Johnson, *Polymer* **13**, 423 (1972).

¹¹J. Blackwell and F. J. Kolpak, *Macromolecules* **8**, 322 (1975).

¹²R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy, Methods and Applications* (Cambridge University Press, Cambridge, 1994).

¹³L. Kuutti, J. Peltonen, J. Pere, and O. Teleman, *J. Microsc.* **178**, 1 (1995).

¹⁴S. J. Hanley, J. Giasson, J.-F. Revol, and D. G. Gray, *Polymer* **33**, 4639 (1992).

¹⁵R. H. Atalla, *The Structure of Cellulose: Characterisation of the solid states*, ACS Symposium, Series 340 (American Chemical Society, Washington, D.C., 1987).

Nanomechanical surface characterization by atomic force acoustic microscopy

U. Rabe, V. Scherer, S. Hirsekorn, and W. Arnold

Fraunhofer Institute for Nondestructive Testing (Izfp), University, D-66123 Saarbrücken, Germany

(Received 12 September 1996; accepted 20 March 1997)

We constructed an atomic force acoustic microscope that enables one to detect out-of-plane and in-plane sample surface vibrations in a wide frequency range from about 100 kHz to several MHz. From the point of view of acoustics, an AFM cantilever is a miniaturized elastic beam that can vibrate in different types of modes as, for example, flexural, torsional, and extensional. We present a method of how to calculate the vibrational behavior of a cantilever with its sensor tip in contact with or near a sample surface, and how to deduce surface properties like the Young's modulus from the vibrations induced in the cantilever. Further, we present single-point measurements in which we compare ultrasonic cantilever vibration spectra where the tip is in contact with surface areas with different stiffnesses to force calibration curves. Finally, we present ultrasonic images taken with our microscope. © 1997 American Vacuum Society. [S0734-211X(97)10604-7]

I. INTRODUCTION

Atomic force microscope (AFM)¹ cantilevers are small elastic beams 100–400 μm in length and a few μm in thickness. They often have an almost rectangular or triangular shape.^{2,3} One end of the beam is fixed to a chip of mm dimensions; the other free end holds the sensor tip. Often the cantilever is regarded as a point mass m^* with a massless spring with stiffness k_c and one resonance frequency $\omega_0 = \sqrt{k_c/m^*}$. However, from the point of view of acoustics the same models can be applied to an AFM cantilever as to a macroscopic beam.^{4–6} If the beam has a homogeneous cross section, its most important free acoustical vibrations are flexural, extensional, and torsional modes.^{7,8} Table I shows the first ten resonance frequencies for a rectangular clamp-free cantilever made of monocrystalline silicon.^{9,10} While the lowest flexural and torsional resonance frequencies of soft cantilevers are often in the kHz range, the longitudinal modes have considerably higher frequencies.

When the sensor tip fixed to the cantilever is brought into contact with a sample surface, flexural or torsional cantilever vibrations can be excited (Fig. 1). A piezoelectric ultrasonic transducer excites longitudinal or shear acoustic waves in the sample under examination. Longitudinal waves generate out-of-plane vibrations $a(t)$ of the sample surface [Fig. 1(a)]. The elastic contact and the adhesion interaction between the tip and the surface force the tip to vibrate. Since the acoustic wavelength is much larger than the tip, one can argue that the forces acting on the tip act as a point force to the free cantilever end and therefore the tip and the free end of the cantilever vibrate at the same amplitude and flexural modes are excited. When a shear wave transducer polarized perpendicularly to the cantilever length axis is coupled to the sample [Fig. 1(b)], primarily the torsional cantilever vibrations will be excited.¹¹ If the sensor tip follows the in-plane surface vibration $a_s(t)$ with a lateral tip amplitude $d_{\text{lat}}(t)$, the corresponding torsion of the cantilever end will be $\theta \approx \tan \theta = d_{\text{lat}}(t)/h$, where h is the tip height. This means that

the sensitivity of the AFM to lateral surface displacements depends not only on the torsional stiffness of the cantilever but also on the length of the sensor tip.

II. IMAGING

The cantilever vibrations forced by a vibrating sample surface can be used for imaging.^{12,13} In our atomic force acoustic microscope, we use a commercial AFM with a beam deflection position sensor. For generating acoustical vibrations we use conventional shear and longitudinal wave transducers with resonance frequencies between ~ 1 and ~ 10 MHz. The sample, with the transducer coupled to it, is imaged in the contact mode. While the topography image is acquired and the feedback loop is activated, the transducer is, for example, excited by tone bursts of 20–50 cycles and several volts of amplitude. The resulting surface displacement generates high-frequency flexural or torsional cantilever vibrations in addition to the low-frequency movements due to topography or friction. These high-frequency pulses are measured by a fast external position detector, a knife-edge detector,¹⁴ that we added to the AFM.¹³ By rotation of the knife edge by 90° either flexural or torsional cantilever vibrations can be detected. At each image point, the measured amplitude or phase of the cantilever vibration signal is fed into the second channel of the AFM, digitized, and displayed as a color coded acoustic image.

As an example of these images, Fig. 2 shows a layered polymer sample that was cut with a microtome; it is composed of two phases with different mechanical properties. Figure 2(a) shows the topography and Fig. 2(b) the acoustical flexural vibration image of the same $15 \times 15 \mu\text{m}^2$ surface area. The excitation signal for the longitudinal wave transducer under the sample was a tone burst at 900 kHz carrier frequency. A rectangular cantilever made of silicon with a spring constant of 22 N/m was used. In Fig. 2(c) the sample topography of a larger $40 \times 40 \mu\text{m}^2$ surface area can be seen

TABLE I. Flexural, torsional, and extensional resonance frequencies of a clamp-free AFM cantilever made of silicon with the dimensions of $260\text{ }\mu\text{m} \times 51\text{ }\mu\text{m} \times 1.5\text{ }\mu\text{m}$ (length \times width \times thickness) and a stiffness k_c of about 0.6 N/m . For the dispersive flexural modes the wave number k_n multiplied by the cantilever length L are also shown.

Mode number n	Flexural waves		Torsional waves	Extensional waves
	$k_n L$	Frequency		
1	1.875	31 kHz	331 kHz	8.19 MHz
2	4.694	192 kHz	992 kHz	24.57 MHz
3	7.885	543 kHz	1.65 MHz	40.95 MHz
4	10.996	1.06 MHz	2.31 MHz	57.32 MHz
5	14.137	1.75 MHz	2.98 MHz	73.70 MHz
6	17.279	2.61 MHz	3.64 MHz	90.08 MHz
7	20.420	3.64 MHz	4.30 MHz	106.46 MHz
8	23.562	4.85 MHz	4.96 MHz	122.84 MHz
9	26.704	6.23 MHz	5.62 MHz	139.21 MHz
10	29.845	7.78 MHz	6.28 MHz	155.59 MHz

and Fig. 2(d) shows the corresponding torsional vibration image. The shear wave transducer was excited with tone bursts of 10 cycles at 1.07 MHz carrier frequency and 5 V amplitude. A triangular cantilever made of Si_3N_4 was used. The stripes that protrude several μm high [material (1)] show a higher signal level than the matrix [material (2)] in the flexural vibration image as well as in the torsional vibration image.

III. FORCED FLEXURAL VIBRATIONS: THEORY

If one assumes that the vibration amplitude and frequency of the sample surface are known, the theoretical problem is to calculate the resulting cantilever vibration signal. The equation of motion for flexural vibrations in a beam with uniform cross section is a differential equation of fourth order.^{7,8} When the sensor tip is not in contact with a sample surface, the situation of a clamp-free cantilever arises which is well known.^{7,8,15} When the sensor tip approaches the sample surface, forces are acting between the end, which was free before, and the sample surface. These forces, which originate from elastic (plastic) deformation and different forces of adhesion, are the subject of investigation because they contain information about, for example, local sample surface elasticity or adhesion. These forces influence cantilever vibration because they transmit acoustic energy from the surface to the tip and they change the boundary conditions at the previously free cantilever.

Although only in a few cases do analytical formulas for tip-sample interaction forces exist, in principle all interaction forces can be attributed to a force curve $F(z)$ which gives the net interaction force at a given tip-sample separation z . $F(z)$ can also display hysteresis. Generally, it is a nonlinear function of z . One possible theoretical approach is to solve the equation of motion numerically taking the full nonlinear interaction force $F(z)$ into account. This was done recently by our group.¹⁶ Figure 3 shows one example where

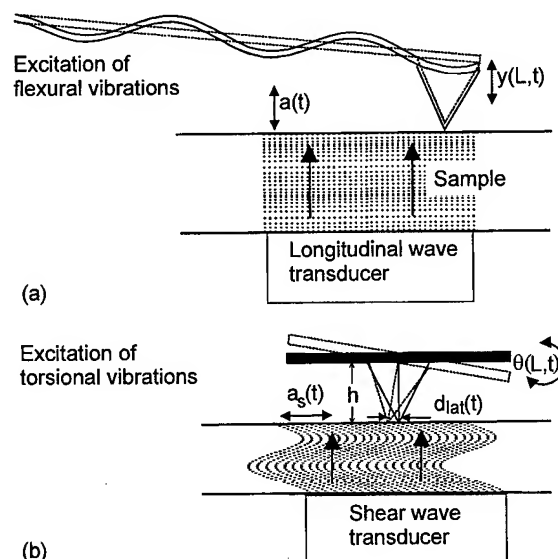


FIG. 1. (a) Excitation of flexural cantilever vibration by a sample insonified with longitudinal waves and (b) shear waves polarized perpendicular to the length axis of the cantilever causing an in-plane vibration of the sample surface, lateral tip movement, and excitation of torsional cantilever vibrations.

the vibration amplitude of a rectangular cantilever in contact with a chromium surface was calculated by analytical derivation of an equation of motion for the surface-coupled end of the cantilever which was then numerically integrated in time. The cantilever data were similar to those used for calculation of the resonance frequencies in Table I. A sample surface vibration at 1.6 MHz carrier frequency with an envelope shown in Fig. 3(a) was the excitation pulse. The force curve $F(z)$ used for the calculation was composed of three different empirical analytical functions and fitted to experimental conditions using a force calibration curve.¹⁷ The resulting cantilever amplitude signal shown in Fig. 3(b) represents a low-frequency deflection away from the sample surface superimposed by a high-frequency vibration. By low-pass and high-pass filtering the two different components of the vibration can be separated and there is very good agreement with experimental results.¹⁸⁻²⁰

Analytical solutions for cantilever vibration can be obtained if vibrations of the sensor tip relative to the sample surface are small. In this case $F(z)$ can be approximated by using its derivative $-k^*$ at the equilibrium tip-sample distance z that was adjusted before the acoustical vibration started. This means that the interaction force is modeled by a linear spring with stiffness k^* fixed between the free cantilever end and the sample surface (Fig. 4). Whereas in the case of a nonlinear interaction force subharmonic or higher harmonic excitation is possible,²¹ in the linear case, when the surface vibrates harmonically out-of-plane with an amplitude $a(t) = a_0 e^{i\omega t}$, the cantilever will vibrate at the same frequency $\omega/2\pi$ as the sample surface. The remaining theoret-

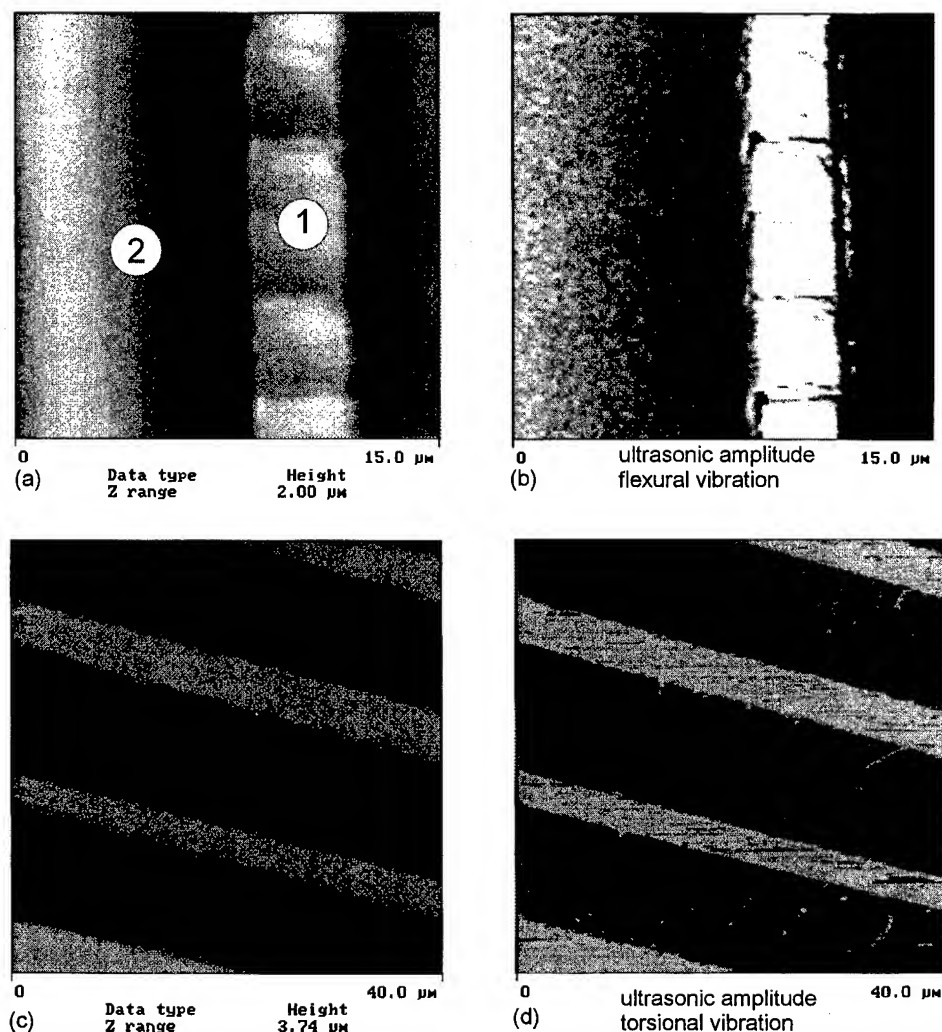


FIG. 2. A polymer sample composed of materials (1) and (2) with different stiffnesses was imaged using flexural cantilever vibrations at a frequency of 900 kHz (b), and using torsional cantilever vibrations at 1.07 MHz (d). (a), (c) Corresponding topography images acquired simultaneously with the acoustical images.

ical problem is, therefore, to calculate the local vibration amplitude and the phase along the cantilever.

In the situation shown in Fig. 4, the relative tip-sample vibration $y(L, t) - a(t)$ gives rise to an elastic force and a damping force summed in F_{ww} :

$$F_{ww} = \left(k^* + \gamma_{ww} \frac{\partial}{\partial t} \right) [y(L, t) - a(t)]$$

$$= (k^* + i\omega \gamma_{ww}) [y(L, t) - a(t)], \quad (1)$$

where $y(L, t)$ describes the vibration of the cantilever at $x = L$ and γ_{ww} is a damping constant that takes into account energy losses due to the tip-sample movement. The static cantilever deflection is chosen to be zero (see Fig. 4). The wave number $k = 2\pi/\lambda$, where λ is the acoustical wavelength, obeys the dispersion relation for flexural vibrations in bars:⁷⁻⁹

$$k = \sqrt{\omega} \cdot \sqrt[4]{\frac{\rho A}{EI}}. \quad (2)$$

Here, E is the modulus of elasticity, ρ is the mass density, A is the cross section, and I is the area moment of inertia of the rectangular cantilever. F_{ww} can now be expressed as

$$F_{ww} = EI \left(\frac{k^*}{EI} + i\gamma_{ww} k^2 \sqrt{\frac{1}{EI\rho A}} \right) [y(L, t) - a(t)]. \quad (3)$$

If this force is added to the appropriate boundary condition, the following complex local vibration amplitude $\tilde{y}(k, x)$ can be calculated:

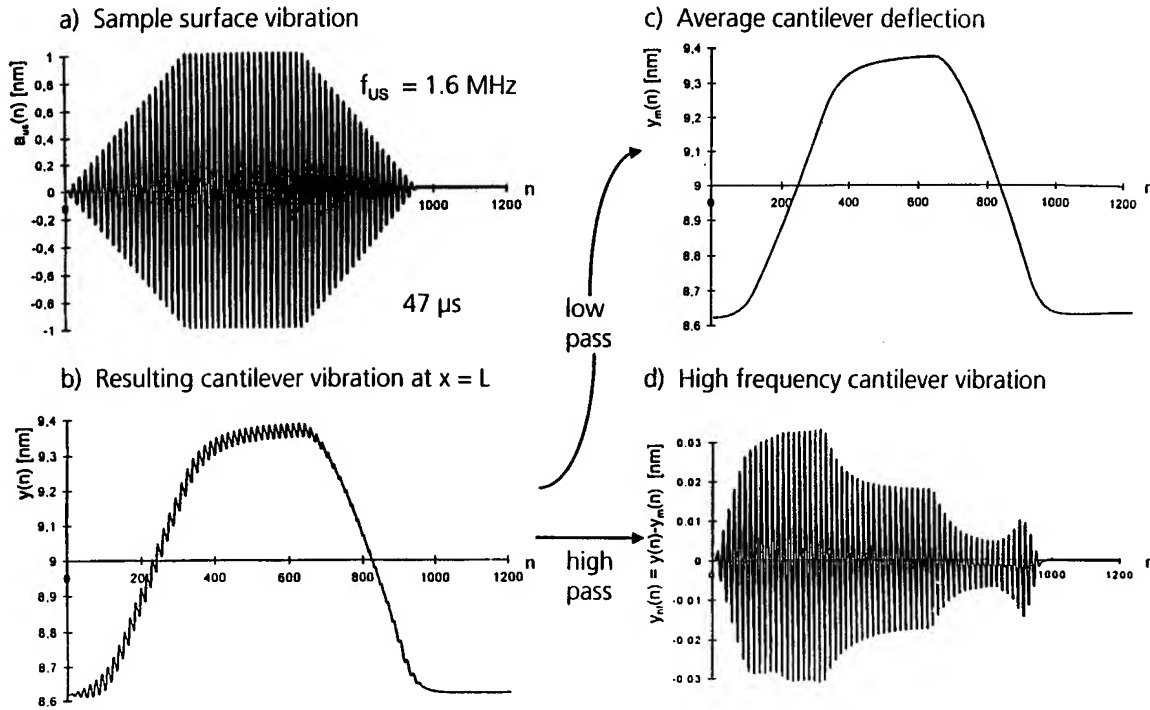


FIG. 3. Result of a numerical calculation of the cantilever vibration using a nonlinear force curve. Vibration amplitude (a) of the sample surface, (b) as an input signal ensuring vibration amplitude at the free end of the cantilever, (c) the low frequency component of (b), and (d) high-frequency component of the cantilever vibration.

$$\tilde{y}(k, x) = \frac{a_0}{2} \left(\frac{-(\sin kL + \sinh kL)(\cos kx - \cosh kx) + (\cos kL + \cosh kL)(\sin kx - \sinh kx)}{k^3} \right. \\ \left. \frac{k^*/EI + i\gamma_{ww}k^2\sqrt{1/EI\rho A}}{(1 + \cos kL \cosh kL) + \cosh kL \sin kL - \sinh kL \cos kL} \right). \quad (4)$$

The local spectral cantilever amplitude $y(\omega, x)$ is obtained by taking the absolute value of the complex amplitude and by using the dispersion relation [Eq. (2)]. Figures 5(a) and 5(b) show two examples where the local vibration amplitude along the cantilever was calculated and plotted for two different contact stiffnesses k^* using Mathematica.²² The vertical axis shows the local vibration amplitude $y(x)$ normalized to the amplitude of the harmonic sample surface vibration a_0 . The frequency of the sample surface vibration is plotted from left to right. The frequency axis is normalized to the lowest flexural vibration frequency $\omega_0 = 2\pi f_0$ of the free cantilever. The third axis is the coordinate along the cantilever.

If k^* is low compared to the cantilever stiffness, the vibration amplitude is maximal at $x=L$. The cantilever behaves like a nearly clamp-free one. If k^* is increased, the resonance is shifted towards higher frequencies. Now the maximum of the vibration amplitude is not at $x=L$ but in the middle of the cantilever, which means that the cantilever now vibrates more like a flexible string fixed at two ends.

Figure 6 shows the normalized cantilever vibration amplitude $y(x)/a_0$ at $x=L$ as a function of the normalized surface vibration frequency ω/ω_0 . The spectral range is larger than

in Fig. 5 because three resonances are now included. Thus one can see again that the resonance frequencies move to higher values if k^* is increased. However, the resonance frequency of every mode remains within a spectral range,

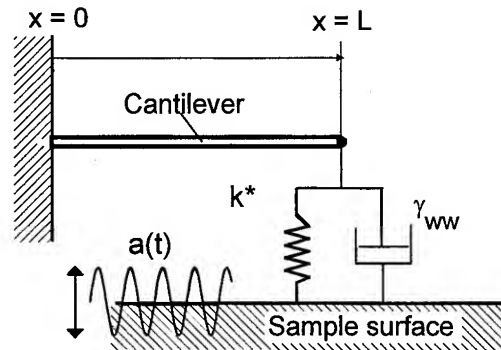


FIG. 4. Linear model for the cantilever vibrating in contact with a sample surface. The clamped end of the cantilever is at $x=0$ and the other end is at $x=L$. The force interaction is approximated by the linear spring k^* . The dashpot γ_{ww} represents damping forces due to tip-sample movement. Cantilever vibrations are excited by the sample surface vibrating out-of-plane with an amplitude $a(t)$.

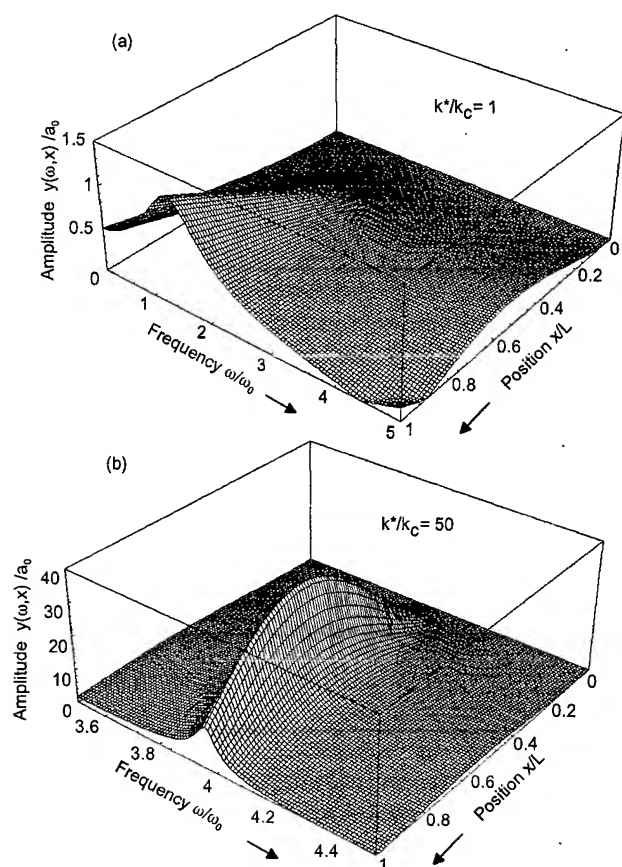


FIG. 5. Three-dimensional plot of the analytical solution for the cantilever vibration obtained with the linear model shown in Fig. 4. (a) $k^*/k_c = 1$, (b) $k^*/k_c = 50$. Please note that the frequency and amplitude scales are different in (a) and (b).

shown by the shaded area. The amplitudes of the resonance peaks increase with increasing k^* . This means the stiffer the spring k^* the better the coupling between the sample surface

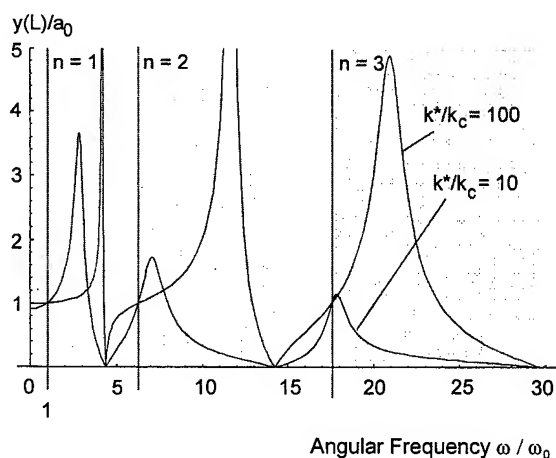


FIG. 6. Vibration amplitude at $x=L$ as a function of the surface vibration frequency ω normalized to the fundamental frequency ω_0 of the free cantilever. The spectral range is chosen so that the first three modes can be seen. Depending on the stiffness k^* every resonance frequency can shift within a fixed spectral range, shown as shaded area.

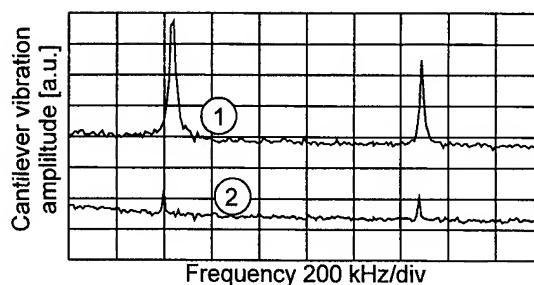


FIG. 7. Experimental spectra measured on the polymer sample shown in Fig. 2. The upper signal was measured on the protruding stripes [material (1)] and the lower signal was measured on the surrounding matrix [material (2)].

and the cantilever that causes a higher cantilever vibration amplitude. Additionally, for the mode $n=1$, mode $n=2$, and the values of k^*/k_c assumed, one observes crossing of the spectra which leads to a contrast reversal²⁰ because not only the amplitude but also the resonance frequency of the peaks depend on k^* .

IV. EXPERIMENTAL DETERMINATION OF THE CONTACT STIFFNESS

Figure 7 shows the cantilever vibration spectra obtained experimentally if the tip is in contact with the polymer sample surface as was seen in Fig. 2. The second and the third resonance frequencies of a rectangular cantilever made of silicon with a spring constant k_c of 22 N/m can be seen. On the striped material (1) which showed higher amplitudes in the acoustical images the resonance maxima have higher amplitude and the resonances are slightly higher in frequency in comparison to those of matrix material (2). The contact stiffness k^* can be calculated from the resonance frequencies⁹ and are listed in Table II.

On samples with a relatively low Young's modulus, like polymer samples, the contact stiffness can also be determined from force calibration curves (Fig. 8). The same cantilever used for the vibration experiment was used here. The vertical axis which displays the cantilever deflection was

TABLE II. Comparison of the values for the contact stiffness k^* of a silicon tip on a polymer sample obtained from force calibration curves and from the resonance frequencies of the cantilever.

Material	k^* calculated from the force curve	Resonance frequencies	k^* calculated from the resonance frequencies
(1) Stripe	21 N/m	655 kHz 1.705 MHz	100–200 N/m
(2) Matrix	4 N/m	615 kHz 1.695 MHz	10–20 N/m

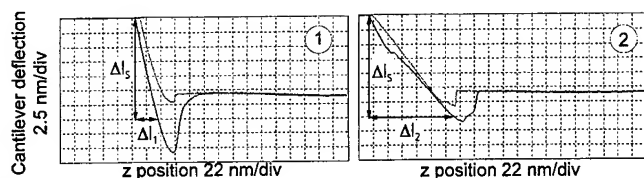


FIG. 8. Force calibration curves from the sample and the same cantilever used in Fig. 7. The two components of the polymer sample are again labeled (1) and (2), the same as image (a) of Fig. 2.

calibrated using a SiC sample with $k_{\text{SiC}}^* \gg k_c$. A difference in slope in the contact region can clearly be seen. The contact stiffness k_i^* of the polymer can then be calculated by considering the cantilever spring k_c and k_i^* as springs in series leading to the formula:

$$k_i^* = \frac{k_c \Delta l_s}{\Delta l_i - \Delta l_s} \quad (5)$$

Here, $i=1,2$ stands for the striped and for the matrix material of the polymer sample, respectively, and Δl_i is the required z displacement to obtain the cantilever deflection Δl_s (see Fig. 8). In accordance with the results from the acoustic experiment, the slope is higher on the striped material (1) which means that the contact stiffness k^* is higher than on the matrix material (2). In Table II the values obtained from the force calibration curves are compared with the values obtained from the resonance frequencies. The fact that k^* calculated from the resonance frequencies is much higher can be attributed to various reasons. The cantilevers are not perfect rectangular beams and damping forces were not considered when evaluating the frequencies. Although care was taken to make the acoustical vibration amplitudes as small as possible, nonlinearities of the force curve might have changed the resonance frequencies. Errors may also arise due to the low spectral resolution that we obtained with the present experimental setup. It is also very possible that polymers show a higher stiffness at higher frequencies. Further experiments and calculations are planned to explain the high k^* values.

From the contact stiffness k^* the Young's modulus of the sample surface can be calculated if, as an approximation, contact models like Hertzian contact mechanics²³ are used. The adhesion forces can be treated like an external force added to the cantilever force. The complete force curve $F(z)$ should be known in principle however. One problem in quantitative determination of surface properties is the unknown shape of the sensor tip. Stable sensor tips with a well-defined radius of curvature and a low cost are still not available.

V. SUMMARY

Apart from the simple contact mode, different AFM imaging techniques were developed where the sample surface or the cantilever is vibrated while the sample surface is scanned. When the AFM cantilever is regarded not only as a

point mass with a massless spring but as a beam that can vibrate in different acoustical modes, a broader theoretical understanding can be obtained for many of these techniques. Furthermore, flexural and torsional vibrations with frequencies up to the MHz range can be used for acoustical imaging. In order to calculate the excitation of flexural cantilever vibrations excited by out-of-plane surface vibrations, the equation of motion of a cantilever coupled to a sample surface with a nonlinear interaction force was solved numerically. With a linear approximation of the interaction force an analytical solution for the cantilever vibration was obtained. From the resonance frequencies of a cantilever vibrating with its sensor tip in contact with a sample surface, the tip-sample interaction stiffness k^* was determined experimentally and compared to values obtained from force calibration curves.

ACKNOWLEDGMENTS

The authors thank Dr. S. Maganov of Digital Instruments, Santa Barbara, CA, for providing the polymer blend sample and Dr. J. Turner for helpful discussions. Financial support by the Volkswagen Foundation and the European Union, Human Capital and Mobility program, is gratefully acknowledged.

- ¹G. Binnig, C. F. Quate, and Ch. Gerber, *Phys. Rev. Lett.* **56**, 930 (1986).
- ²NanoSensors, Dr. Olaf Wolter GmbH, Wacholderweg 8, D-71134 Aidlingen, Germany.
- ³R. Albrecht and C. F. Quate, *J. Vac. Sci. Technol. A* **6**, 271 (1988).
- ⁴D. Sarid, *Scanning Force Microscopy* (Oxford University Press, Oxford, 1991).
- ⁵G. Y. Chen, R. J. Warmack, T. Thundat, D. P. Allison, and A. Huang, *Rev. Sci. Instrum.* **65**, 2532 (1994).
- ⁶D. F. L. Jenkins, M. J. Cunningham, W. W. Clegg, and M. M. Bakush, *Meas. Sci. Technol.* **6**, 160 (1995).
- ⁷W. F. Stokey, in *Shock and Vibration Handbook*, edited by C. M. Harris and C. E. Crede (McGraw-Hill, New York, 1976), pp. 7-1 ff.
- ⁸P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968).
- ⁹U. Rabe, K. Janser, and W. Arnold, *Rev. Sci. Instrum.* **67**, 3281 (1996).
- ¹⁰U. Rabe, Ph.D. thesis, Universität des Saarlandes and Fraunhofer Institute for Nondestructive Testing, Saarbrücken, 1996 (unpublished).
- ¹¹V. Scherer and W. Arnold, *Micro/Nanotribology and its Applications*, edited by B. Bushan (Kluwer Academic, Dordrecht, 1997), pp. 225-231.
- ¹²P. Maivald, H.-J. Butt, S. A. C. Gould, C. B. Prater, B. Drake, J. A. Gurley, V. B. Elings, and P. K. Hansma, *Nanotechnology* **2**, 103 (1991).
- ¹³U. Rabe and W. Arnold, *Appl. Phys. Lett.* **64**, 1493 (1994).
- ¹⁴J. W. Wagner, in *Physical Acoustics* edited by R. N. Thurston and A. D. Pierce (Academic, London 1990), Vol. XIX, p. 201.
- ¹⁵S. Timoshenko, D. H. Young, and W. Weaver, *Vibration Problems in Engineering* (Wiley, New York, 1974).
- ¹⁶S. Hirsekorn, U. Rabe, and W. Arnold, *Nanotechnology* **8**, 57 (1997).
- ¹⁷S. Hirsekorn, U. Rabe, and W. Arnold, *Europhys. News* **27**, 93 (1996).
- ¹⁸U. Rabe, K. Janser, and W. Arnold, *Acoustical Imaging*, edited by P. Tortoli and L. Masotti (Plenum, New York, 1996), Vol. 22, pp. 669-676.
- ¹⁹O. Kolosov and K. Yamanaka, *Jpn. J. Appl. Phys.* **1** **32**, 22 (1993).
- ²⁰N. A. Burnham, A. J. Kulik, G. Gremaud, P.-J. Gallo, and F. Oulevey, *J. Vac. Sci. Technol. B* **14**, 794 (1996).
- ²¹N. A. Burnham, A. J. Kulik, G. Gremaud, and G. A. D. Briggs, *Phys. Rev. Lett.* **74**, 5092 (1995).
- ²²S. Wolfram, *Mathematica*, A System for Doing Mathematics by Computer, Wolfram Research Inc., 100 Trade Center Drive, Champaign, IL 61820-7237.
- ²³K. L. Johnson, *Contact Mechanics* (Cambridge University Press, Cambridge, 1985).

Detection mechanism of an optical evanescent field using a noncontact mode atomic force microscope with a frequency modulation detection method

Masayuki Abe^{a)} and Takayuki Uchihashi

Department of Physics, Faculty of Science, Hiroshima University, 1-3-1 Kagamiyama, Higashi-hiroshima 739, Japan

Masahiro Ohta, Hitoshi Ueyama, Yasuhiro Sugawara, and Seizo Morita

Department of Electronic Engineering, Faculty of Engineering, Osaka University, 2-1 Yamadaoka, Suita 565, Japan

(Received 8 September 1996; accepted 19 May 1997)

By using the noncontact atomic force microscope with a frequency modulation detection method, the force gradient induced by the optical evanescent field was detected in a high vacuum. We succeeded in measuring the exponential distance dependence of the force gradient induced by the optical evanescent field. Furthermore, we investigated the incident beam intensity and bias voltage dependence of the force gradient induced by the optical evanescent field. We confirmed that the detection mechanism is not photothermal effect but the surface photovoltage effect. © 1997 American Vacuum Society. [S0734-211X(97)13504-1]

I. INTRODUCTION

Various types of scanning near-field optical microscopes (SNOMs)¹ with subwavelength resolution have been developed as alternatives to a conventional optical microscope. SNOM is one typical application using evanescent fields. As an usual technique in SNOM, the (optical) evanescent field has been usually detected by using a tapered optical fiber with a subwavelength sized aperture^{2,3} or small scattering tip.^{4,5}

Recently, as an alternative method to detect the evanescent field, detection of image force between a semiconductor atomic force microscope (AFM) tip and a glass surface was proposed.⁶ In that experiment, modulation frequency of an incident beam was set to the resonance frequency of the cantilever to make use of the resonance enhancement, and the force induced by the evanescent field was estimated from the cantilever deflection. It was clarified that the origin of force is based on the phenomenon called surface photovoltage (SPV). Furthermore, the performance of this novel method as near-field optical microscopy was preliminarily demonstrated by imaging the standing evanescent wave in a moderate vacuum or in ambient.⁷

Recently, by using the noncontact mode AFM with the frequency modulation (FM) detection method,⁸⁻¹⁰ we measured the distance and polarization dependence of the force gradient induced by the evanescent field.¹¹ In this novel method, there are two advantages: One is extreme high sensitivity for force detection, because of the sensitivity enhancement due to a very high Q value of the cantilever in a vacuum.¹² The other is the possibility of accurate estimation of the local force gradient, because capillary forces caused by the water film on the sample surface can be completely removed in a high vacuum ($\sim 10^6$ Torr).

In this article, we investigate the force gradient induced

by the evanescent field. Further, from the measurements of the incident beam intensity and bias voltage dependence, we also examine the SPV model for the force mechanism in SNOM.

II. EXPERIMENT

Figure 1 shows the schematic diagram of the noncontact mode AFM to detect the evanescent field. The evanescent field was generated by total internal reflection of a laser beam at the interface between a glass prism and a vacuum. The force gradient F' acting on the silicon AFM tip was measured using the FM detection method. In this method, the piezoelectric tube scanner was used to vibrate the cantilever at its mechanical resonance frequency, ν_0 . The positive feedback system with an automatic gain control (AGC) circuit was used to maintain the vibration amplitude of the cantilever at a constant level. The frequency shift $\Delta\nu$ of the cantilever was detected by a tunable analog FM demodulator. Here, the force gradient F' acting on the AFM tip is approximately given by the following equation⁸:

$$F' \cong 2k \frac{\Delta\nu}{\nu_0}, \quad (1)$$

where k is the spring constant of the cantilever. When the AFM tip approaches the glass prism surface, both van der Waals attractive force and the image force due to the evanescent field exerts on the AFM tip. Here, the incident beam was modulated at a frequency ν_{chop} by the optical chopper, and the force due to the evanescent field was detected as an ac component of the force gradient at a frequency ν_{chop} . On the other hand, the van der Waals force was measured from the dc component of the force gradient.

The detection of the evanescent field using the AFM tip was performed in a high vacuum (10^{-6} – 10^{-4} Torr). A triangle-shaped prism was used to generate the evanescent field. Light beam from laser diode ($\lambda=690$ nm, 30 mW), He-Ne laser ($\lambda=633$ nm, 7 mW), or Ar laser ($\lambda=488$ nm,

^{a)}Electronic mail: masa@ele.eng.osaka-u.ac.jp

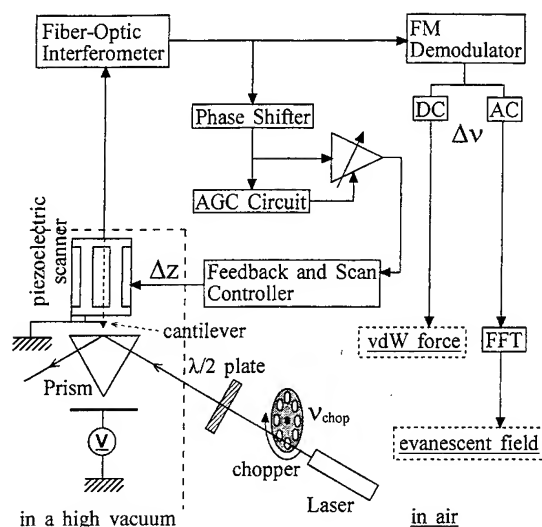


FIG. 1. Schematic diagram of noncontact mode AFM with the FM detection to detect the optical evanescent field.

25 mW) was incident through a viewpoint of a vacuum chamber. Incident angle with the glass prism surface was 60° .

Sb-doped conductive silicon cantilevers with a sharp tip were used as a force sensor. When the attractive force gradient F' acting on the AFM tip exceeds the spring constant k of the cantilever, the cantilever jumps into the sample surface. So, for a force sensor, stiff silicon cantilevers were used to suppress the jump of the cantilever and to avoid the crush of the initial sharp tip.¹⁰ The resistivity of the cantilever was $\rho = 0.01\text{--}0.02\ \Omega\text{ cm}$. The spring constant k and the resonance frequency ν_0 of the cantilever were $k = 35\text{--}57\text{ N/m}$ and $\nu_0 = 150\text{--}170\text{ kHz}$, respectively. The radius of curvature, the apex angle, and the height of the tip were $5\text{--}15\text{ nm}$ and $40\text{--}45^\circ$ and $10\text{--}15\ \mu\text{m}$, respectively.

III. RESULTS AND DISCUSSIONS

When the AFM tip approaches the prism surface without irradiation of the incident beam, the dc component of the force gradient F' due to the van der Waals force remained zero for the AFM tip to prism surface distance larger than $2\text{--}3\text{ nm}$, and also the ac component was at noise level. However, with irradiation of the incident beam, the ac component of the force gradient F' was observed even in the far region more than $2\text{--}3\text{ nm}$. Here, we define that $z = 0$ is the distance where the AFM tip begins to be in contact with the surface; at $z = 0$, the positive feedback system with the AGC circuit for vibrating the cantilever starts to be unstable. Figure 2 shows the fast Fourier transform (FFT) spectrum of the FM demodulator when He-Ne laser was incident at a modulation frequency $\nu_{\text{chop}} = 100\text{ Hz}$. We can see that the force gradient acting on the AFM tip was modulated at frequencies of 100, 300, and 500 Hz by the incident beam. Peaks at 300 and 500 Hz are harmonics induced by the harmonics of the incident beam because the mechanical chopper was used to modulate the incident beam.

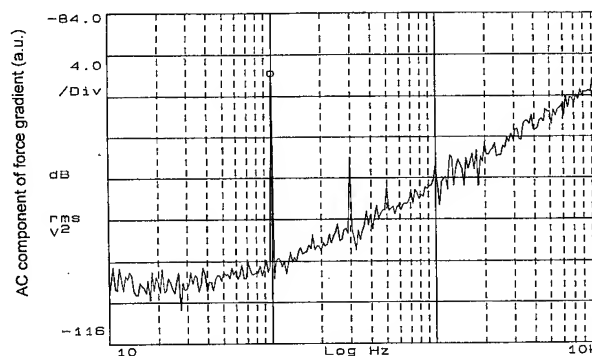


FIG. 2. FFT spectrum of the FM demodulator output when He-Ne laser without focusing was incident at a modulation frequency $\nu_{\text{chop}} = 100\text{ Hz}$. Using a mechanical chopper to modulate the incident beam, it resulted in peaks at frequencies of 100, 300, and 500 Hz. These peaks indicate the force gradient induced by the evanescent field.

Figure 3 shows the force gradient F' distance curve induced by the evanescent field as a function of the distance between the prism and the AFM tip. Incident beam from laser diode at modulation frequency 1 kHz was used. Spectrum analyzer with analog output was used to measure the force gradient F' . Bias voltage at -10 V was applied to the prism to enhance the signal. As the tip approaches the prism surface, the force gradient F' increased. From this experiment, we confirmed the exponential distance dependence of the force gradient.

There are two possibilities for the light-induced force interaction mechanism between the incident beam and the AFM tip. One is the thermal expansion effect for the prism surface and/or the tip apex due to the photothermal effect of

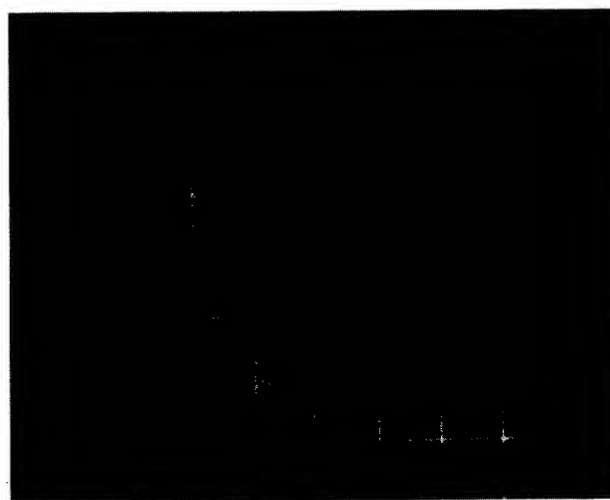


FIG. 3. Force gradient F' distance curve induced by the evanescent field. The measurement took about 5 s. The horizontal axis corresponds to the distance between the prism and the AFM tip (10 nm/div). Vertical axis corresponds to the force gradient F' (arbitrary unit) induced by the evanescent field. Laser diode was used for incident beam modulated at a frequency of 1 kHz . Spectrum analyzer with analog output was used to measure the force gradient induced by the evanescent field. Resolution bandwidth and video bandwidth of the spectrum analyzer were 300 and 10 Hz , respectively. -10 V was applied to the prism to enhance the signal.

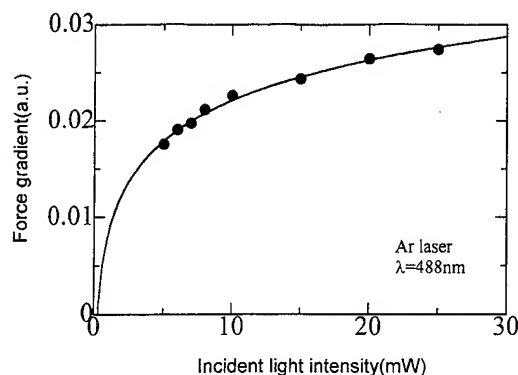


FIG. 4. Incident beam intensity dependence of the force gradient induced by the evanescent field. Ar laser without focusing was used for incident beam. Horizontal axis corresponds to the incident beam intensity. Vertical axis corresponds to the force gradient F' . Circles (●) indicate the measured force gradient. Solid line indicates the fitted result of measured data as a function $a + b \ln(I)$ where I is incident beam intensity, and a and b are fitting parameters.

the incident beam. The other is the SPV effect induced by the band bending on the silicon AFM tip surface suggested by Mertz *et al.*⁶

In order to confirm the SPV model, the incident beam intensity and bias voltage dependence of the force gradient F' were investigated. In the SPV model, force induced by the evanescent field became saturated with the increase of incident beam intensity,⁶ and varies linearly with bias voltage.⁷ Figure 4 shows the measured incident beam intensity dependence of the force gradient F' . Circles (●) indicate the measured force gradient. Solid lines indicate the calculated result by least square method. The force gradient F' seems to be saturated with the increase of incident beam intensity. This tendency agrees well with the SPV model. Figure 5 shows the measured bias voltage dependence of the force gradient F' induced by the evanescent field. Circles (●) indicate the measured force gradient. Dotted lines indicate the noise level of the force gradient. The force gradient F' varied almost linearly with bias voltage from -10 to 0 V, and agrees well with the SPV model in this regime. However, the force gradient F' could not be detected at applied voltage larger than $+3$ V. In the experiments of Figs. 4 and 5, we have obtained the same results in different glass prisms. Tentatively, we attributed the reason why the force interaction was not observed, to the effect due to the surface state density of the semiconductor AFM tip. That is, in the Hipp *et al.*'s experiment,⁷ surface state density of the semiconductor AFM tip with oxide layer was high and the Fermi level was pinned. In their experiment, bias voltage was applied to the tip through an ohmic contact. Hence the force induced by the evanescent field varied linearly with the applied bias voltage. However, in the present experiment, the oxide layer on the AFM tip seems to be removed due to the slight contact with the prism surface in a high vacuum, and surface state density of the semiconductor AFM tip seems to be low and the Fermi level was unpinned. In this condition, band bending and hence barrier capacitance are dependent on

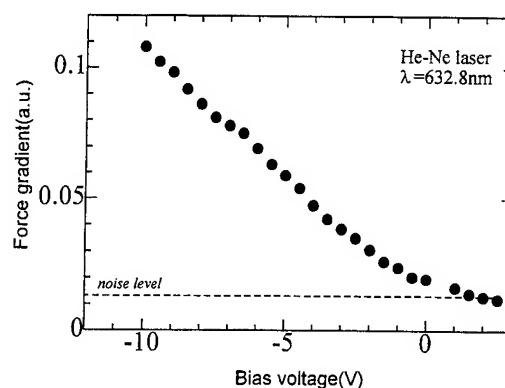


FIG. 5. Bias voltage dependence of the force gradient induced by the evanescent field. He-Ne laser without focusing was used for incident beam. Horizontal axis corresponds to the bias voltage applied to the sample as shown in Fig. 1. Vertical axis corresponds to the force gradient F' . Circles (●) indicate the measured force gradient. The measured force gradient decreased linearly with bias voltage from -10 to 0 V. This result agreed with the SPV model. However, force gradient could not be detected at applied voltage larger than $+3$ V.

the bias voltage. Barrier capacitance is increased with decrease of the bias voltage in the negative bias region, while that is almost constant in the positive bias region. As a result, in the positive bias region, although the semiconductor AFM tip is positioned in the near-field region and electron-hole pairs are created, SPV will be unchanged. Namely, the force gradient seems not to be detected at the voltage larger than $+3$ V.

IV. CONCLUSION

By using the noncontact mode AFM with FM detection method in a high vacuum, the force gradient induced by the evanescent field was detected. There are two advantages of this method. One is high sensitivity as a force sensor. The other is that the local force gradient can be estimated without influence of capillary force due to water film on the prism surface. Using this method, we measured the force gradient distance curve induced by the evanescent field. We measured the incident beam intensity and bias voltage dependences of the force gradient induced by the evanescent field. As a result, we confirmed the SPV model suggested by Mertz *et al.*⁶ In the future, the SPV model needs to be modified to explain the results of the distance and polarization dependence of the force gradient.¹¹

ACKNOWLEDGMENTS

Parts of this work were supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports, and Culture of Japan, Foundation Advanced Technology Institute, and Inoue Foundation for Science.

¹D. Courjon, F. Baida, C. Bainier, D. Van Labeke, and D. Barchiesi, *Photons and Local Probes*, edited by O. Marti and R. Möller (Kluwer Academic, Dordrecht, 1993).

²R. C. Reddick, R. J. Warmack, and T. L. Ferrell, *Phys. Rev. B* **33**, 767 (1989).

- ³M. Ohtsu, J. Lightwave Technol. **13**, 1200 (1995).
- ⁴F. Zenhausern, Y. Martin, and H. K. Wickramasinghe, Science **269**, 1083 (1995).
- ⁵K. Fukuzawa, Y. Tanaka, S. Akamine, H. Kuwano, and H. Yamada, J. Appl. Phys. **78**, 7376 (1995).
- ⁶J. Mertz, M. Hipp, J. Mlynek, and O. Marti, Appl. Phys. Lett. **64**, 2338 (1994).
- ⁷M. Hipp, J. Mertz, J. Mlynek, and O. Marti, *Photons and Local Probes*, edited by O. Marti and R. Möller (Kluwer Academic, Dordrecht, 1993), NATO ASI Series Vol. 300, p. 109.
- ⁸H. Ueyama, M. Ohta, Y. Sugawara, and S. Morita, Jpn. J. Appl. Phys. **34**, 1086 (1995).
- ⁹Y. Sugawara, M. Ohta, H. Ueyama, and S. Morita, Science **270**, 1646 (1995).
- ¹⁰Y. Sugawara, M. Ohta, H. Ueyama, F. Osaka, S. Ohkouchi, M. Suzuki, S. Mishima, and S. Morita, J. Vac. Sci. Technol. B **14**, 953 (1996).
- ¹¹M. Abe, T. Uchihashi, M. Ohta, H. Ueyama, Y. Sugawara, and S. Morita, Opt. Rev. **4**, 232 (1997).
- ¹²T. R. Albrecht, P. Grütter, D. Horne, and D. Rugar, J. Appl. Phys. **69**, 688 (1994).

Fabrication of probe tips for reflection scanning near-field optical microscopes: Chemical etching and heating-pulling methods

Mufei Xiao,^{a)} Jesús Nieto, and Roberto Machorro

Laboratorio de Ensenada, Instituto de Física, Universidad Nacional Autónoma de México Apartado Postal 2681, CP 22800 Ensenada, Baja California, México

Jesús Siqueiros^{b)} and Hector Escamilla

Centro de Investigación Científica y Educación Superior de Ensenada Apartado Postal 2732, CP 22800, Ensenada, Baja California, México

(Received 12 September 1996; accepted 31 March 1997)

Work is described on the fabrication of uncoated probe tips for reflection scanning near-field optical microscopes. A chemical etching process with a rotator, and a simple device to make the tips by a heating-pulling method without a heating laser are introduced. Both techniques allow some manipulation of the tips. Merits and demerits of these two methods are discussed. Finally, it is suggested that observation of the far field scattering pattern of the tips can be used to check the usefulness of the tips. © 1997 American Vacuum Society. [S0734-211X(97)04004-3]

I. INTRODUCTION

In scanning near field optical microscopy (SNOM),¹ the mainstream is still to use an optical fiber tip scanning above the sample surface to pick up the optical near field.² In the microscopy, the signals received by the probe tip contain both the evanescent field and the propagating waves. To ensure a high proportion of the evanescent field collection, the probe tip has to be extremely small and close to the surface. The smallness of the probe tip enables the field components with higher spatial frequencies to be collected. The probe tips are commonly made of tapered optical fiber with end sizes in the range of <100 nm. In the so called transmission mode SNOM (TSNOM),³ the tips are coated with metal to form a 20–50 nm aperture at the very end of the tip, while uncoated tips are widely used in the reflection mode SNOM (RSNOM).^{4–7} Both single mode and multimode optical fibers have been in use.

There have been two methods to fabricate the fiber tips, namely the chemical etching^{2,8} and heating-pulling process.^{3,9,10} The chemical etching method is easy to implement with low cost and is still in use. The smallest tips can be made by carefully choosing the chemical solvent⁸ and high optical transmission efficiency is realizable.¹¹ However, the acids used to etch the fibers are highly poisonous, which limits the use of the method. In SNOM, the widely used method to form tips is the heating-pulling process.^{3,9} The pulling is usually realized by a commercial puller and is controlled by a computer. The heating is performed by a laser (CO₂ or other type) with a typical laser power 10–30 W and beam spot sizes of 2–3 mm. The laser may also be controlled by the computer. A sophisticatedly designed heating-pulling setup provides tips suitable for most of the applications in SNOM. However, an obvious disadvantage of

the setup is that the high cost may prevent small research groups from being equipped with it.

In the present article, we report our work on both chemical etching and heating-pulling processes. We have developed a controllable etching process. In the process, we rotate the etching acid, which brings about several advantages: faster etching, smoother tip surface, better symmetry, and a controllable length of the tip inside the acid. In order to better control the tip size, we used a multistep etching scheme. The tips are washed with purified water and checked by optical microscopy every 5–10 min so that the fabrication is more accurately controlled.

Despite the simplicity of the chemical etching process, the setup is not desired in the work place mainly because of the hazardous acid. The chemical etching takes a long time, and also, the remains of acid on the tip surface are difficult to be washed away completely. In order to realize a fast, reliable and *in situ* tip forming, we have designed a simple heating-pulling device for making fiber tips. A resistance coil replaces the laser, while the puller is driven by a mechanical spring. Comparing with the common laser-pull systems, the cost of our setup is actually negligible and the tip fabrication can be realized in any place, even with the fiber aligned on the microscope. Despite its simplicity, the device is able to produce equally good tips in comparison with the conventional laser-pull systems.

The above two techniques are comparatively discussed in this article. Advantages and technical restrictions are outlined. Finally, we discuss how to check the tips before use in microscopes. We propose to measure the far field scattering patterns. Based on the far field pattern, the usefulness of the tips may be estimated, especially for RSNOM applications.

The organization of the article is as follows. In Sec. II, we describe the novel features of the chemical etching process. In Sec. III, we describe the setup and process for heating-pulling. In Sec. IV, we take notes on advantages and general technical restrictions of both processes. We discuss the characterization of the tips by measuring the far field patterns.

^{a)}Mailing address: Institute of Physics—UNAM, P. O. Box 439036, San Ysidro, California 92143; Electronic mail: mufei@ifisicaen.unam.mx

^{b)}On Sabbatical leave from IFUNAM, Lab. Ensenada.

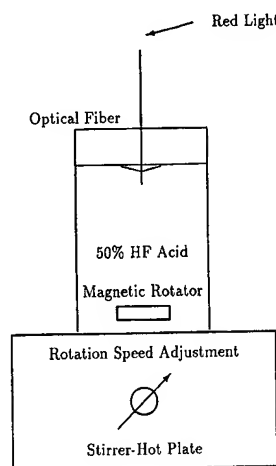


FIG. 1. Schematic of the etching setup with a magnetic rotator.

II. THE ETCHING

In this section, we present technical notes on our chemical etching process. Chemical etching has for a long time been used to form SNOM tips. The technique was used in the first SNOM experiments using optical fiber tips.² Since then many groups have used and improved the process. In Ref. 8, some important aspects in the chemical etching process were discussed, and in particular, it was pointed out that selective etching may be realized with a composed etching solution and it would be helpful to etch the fiber in two steps with different chemical solutions. In the present article, we emphasize some novel features we have introduced in our experiments. We have previously described in general the etching process with some manipulation of the tip shape.¹² Progress has been made recently by employing a rotator in the etching solution and a multistep etching scheme.

The use of a rotator in the etching is schematically shown in Fig. 1. We have used earlier in our experiments a rotating motor to hold the acid container. Recently, we have replaced the motor by a rotator, which is easier to operate with. The rotator is a magnetic cylinder sank in the etching solution. The rotator is driven magnetically by a stirrer-hot plate (VWR502 by DYLA-DUAL). The rotation speed can be controlled by the plate. The rotation results in eddy flow of the liquid surface. In the center of the liquid surface, the surface height goes down with the increase of the rotation speed (see in Fig. 1). The higher the speed, the lower the surface height. The optical fiber is introduced vertically into the acid at the center of the liquid surface. During the process, a red light is coupled into the fiber from the other end (or the fiber is already aligned in the microscope). The visible light spot at the fiber end is helpful to determine the immersion length of the tip. The light spot is also important to monitor the process, which will be discussed in Sec. IV in this article. When the fiber end is immersed into the liquid, due to surface tension, the liquid tends to creep up along the fiber. This problem can be partially avoided by the rotation of the liquid. Also, the fine adjustment of the immersion length may be realized by changing the rotation speed. Gen-

erally speaking, by using the rotation technique, we have improved the etching in the following aspects. First of all, the etching time is shortened considerably. To form a tip with end size ~ 50 nm, the etching can be done in 20–30 min, instead of 120–240 min if the liquid is kept static. Second, the fabricated tips come out smoother and more symmetric compared to the static etching. It is worth mentioning that we have checked a few fabricated tips with a scanning electron microscope (SEM) and we have not found tips with end sizes smaller than ~ 20 nm, which seemed to be the size limit of our etching process. Note that in Ref. 8, tips with end sizes of a few nanometers were reported. We stress that in the RSNOM configuration, if the fiber is used both as the light emitter and the collector, one has to ensure sufficient incident light.⁵ Extremely small tips may not be useful in RSNOM (see discussions in Sec. IV).

A difficulty in the etching was that the etching time appears not to be an accurate controllable parameter for the process even with all other conditions fixed in the experiments. In a certain range of time, the formed tips vary in size and shape. In light of this, we have designed a practical multistep etching scheme. During the etching, the tip is washed in purified water every 5–10 min and, from time to time, checked with a $\times 25$ optical microscope, so that promising tips may be identified. The change of the scattering pattern, observed in a dark room or photographed with the microscope, is a good indication of the usefulness of the tips. Note that SEM photography may allow better evaluation of the tip, but the method is unfortunately destructive and more importantly, the relation between the tip shape and the imaging quality remains, to a large extent, an open question both in theory and in experiment.

In the experiments, we have used 50% HF acid and Newport F-MSD fibers (core $50\text{ }\mu\text{m}$ and cladding $125\text{ }\mu\text{m}$). Using the etching process, we have fabricated tips ranging from 20 to 200 nm in size. A photograph of a fabricated tip is presented in Sec. IV, where we also discuss the estimated usefulness of the tips by observing the far field scattering pattern.

III. THE HEATING PULLING

As already mentioned in the introduction, despite the simplicity of using the above described etching process to make the tips particularly useful for RSNOM, the etching acid is not desirable in the laboratory, especially if the tip has to be made on the spot. Motivated by the absence of the high power laser and the multistage puller in our laboratory, we have designed and fabricated a simple heating-pulling device for making fiber tips. We have been able to substitute the laser and multistage puller that have been widely used in the field.

In Fig. 2, the device is depicted. The fiber end is firmly held by two identical holders made of teflon. The holder can move smoothly inside a fixed aluminum tube. The two holders are pulled by a spring through a fishing thread of nylon ($\phi=0.5$ mm). The spring is fixed at one end, and the fishing thread is fixed at the free end of the spring. The free end of

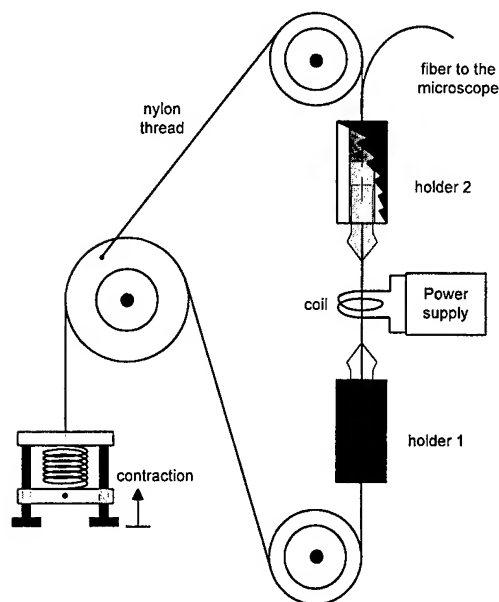


FIG. 2. Schematic of the heating-pulling setup.

the spring can be contracted to provide the pulling force. The contracted length controls the initial pulling force. A nickel-chrome wire ($\phi=0.5$ mm) forms a one turn coil of diameter 2.5 mm, which is driven by a power supply. The temperature distribution in the center of the coil is, therefore, a simple function of the driving current.

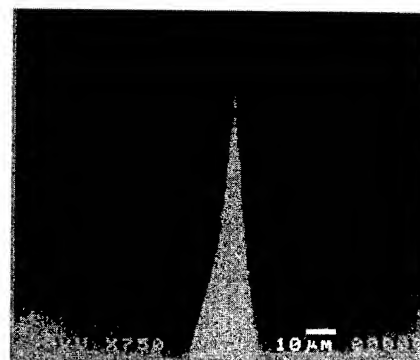
In the device, the pulling is driven by a mechanical spring. We have carefully chosen the spring (1.6 cm long, 1.5 cm diameter, five turns of steel wire with $\phi=0.9$ mm), and calibrated the contraction range, so that the pulling is performed by the well known spring potential which decreases gradually during the thinning process.

Let us now describe the experimental steps. First of all, one fixes the fiber and contracts the spring. One then increases the current to set the highest temperature, which takes ~ 1 min to have the temperature established. For most fibers, 1100–1200 °C is sufficient. The fiber breaks apart eventually after about 2–3 min. The setup has to be mounted vertically to avoid forming asymmetric tips. We have checked the tips with an optical microscope of $\times 25$ for a rough estimation and with a SEM for final characterizing. The shape and size can be varied by changing the contraction length of the spring and the current of the coil. For a specific fiber, once the setup is adjusted for making desirable tips, the fabrication is reliable.

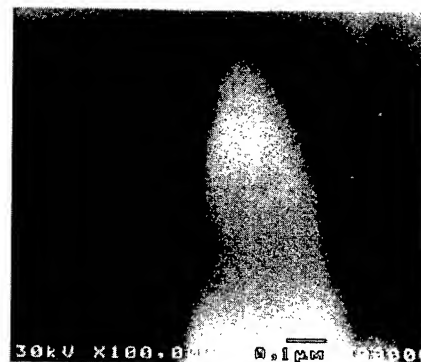
A tip formed by the device is shown in Fig. 3. The optical fiber was F-MSD from Newport with a 50 μm core and a 125 μm cladding. The tip size is well within 50 nm [see Fig. 3(b)].

IV. SOME DISCUSSIONS

In this section, we first present a comparative analysis on the advantages and disadvantages of the above two methods. As far as uncoated tips in RSNOM are concerned, it seems desirable that the very end of the fiber tip acts like a pointlike



(a)



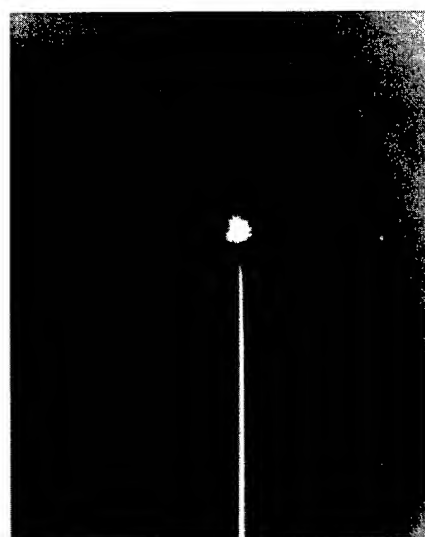
(b)

FIG. 3. SEM photos of a pulled tip, (a) 10 μm bar; (b) 0.1 μm bar. The dots are guide of eyes.

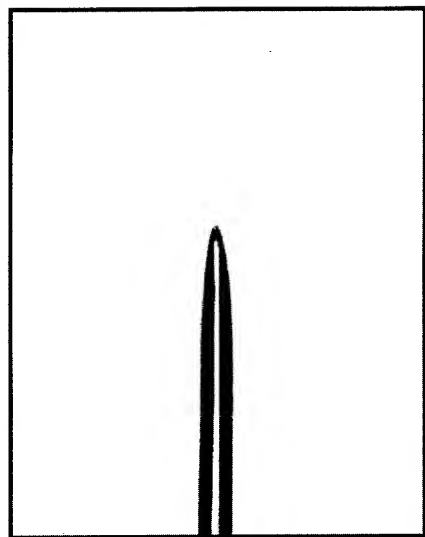
emitter⁵ and the light leakage before the very end should be avoided. For a workable fiber tip, one is able to experimentally determine such a scattering center.^{5,12} Our experiences with the above two fabrication processes have shown that the chemical etching makes tips more like a point scatter. This is because the etching can make tips with larger taper angles (typically 20°–50°) and the tip surface is smoother in comparison to the pulled tips. In the light of the above analysis, we particularly recommend our etching process for making interactive mode fiber tips that are often used in RSNOM. The developed heating-pulling process can also make workable tips but the leakage of the light through the sidewall of the tips may complicate the received signals. The pulled tips, which have usually small taper angles, are however suitable to be coated and used in other configurations in SNOM.

The primary disadvantage of using the etching technique is the harmful environment it creates, and also the longer etching time. The choice between the two methods should be made in consistence with specific experimental conditions.

In the following, we discuss how to estimate the usefulness of the fabricated tips. Apparently, one may want to know the accurate physical shape of the tip, which may be realized with high resolution microscopes, such as a SEM. However, the microscopes may be either destructive or too remote for *in situ* tip checking. Even if one can have the accurate photograph of the tips without destroying them, one still does not know whether the tip can have a good imaging



(a)



(b)

FIG. 4. Photos of an etched fiber tip with (a) and without (b) light scattering. The diameter of the tip is 50 μm .

quality in RSNOM applications. So far, the relations between the tip shape and the resolution of RSNOM are unknown both theoretically and experimentally. Some preliminary experiments results in SNOM configuration to determine the polarization state were reported in Ref. 13. Theoretically, three dimensional modeling to relate the tip shape and light scattering to the near field imaging quality has not been reported. To determine the usefulness of the tips, one may use the tip in RSNOM with known samples. However, the estimation is rough because there are many parameters involved in the interactive RSNOM configuration.¹⁴ Also, it should be convenient to have the estimation before the use of the tips. For this purpose, we have been attempting to relate the far field scattering to the imaging quality. The far field scattering is easy to observe

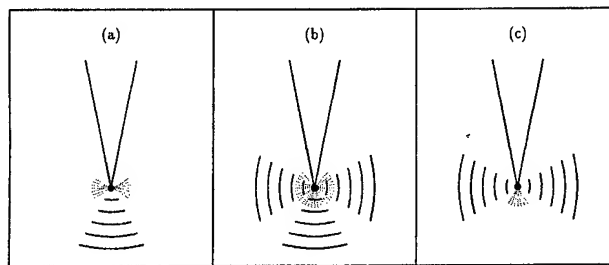


FIG. 5. Schematic of far field scattering pattern of a tapered fiber tip, vertical (a), circle (b), and parallel (c).

and record, which can be done during the fabrication process. We keep the light coupled into the fiber during the etching (or after the pulling for pulled tips). Different scattering patterns may be observed in a dark environment, ensuring that the scattering occurs only at the very end of the tip (pointlike scatterer). Identified tips are further checked with the $\times 25$ microscope, and if necessary, recorded by camera. In Fig. 4, a fabricated tip is photographed under the $\times 25$ microscope with [Fig. 4(a)] and without [Fig. 4(b)] the light scattering. From the photographs, one realizes that though the very end of the tip is not completely resolved by the microscope, the light scattering pattern has been recorded, and looks like a circle. In summary, the scattering pattern of a tip acting as a pointlike scatterer can be identified by simple observation.

The remaining question is how to determine which kind of the scattering pattern is desired in the RSNOM experiments. A preliminary discussion of this problem may be found in Ref. 12. Generally speaking, three typical patterns may be easily identified, which are schematically drawn in Fig. 5. In Fig. 5(a), the light is concentrated toward the sample surface, while in Fig. 5(c) the light goes along the surface, and in Fig. 5(b) the scattering is a circle (as in the case of Fig. 4). Supposing the tips are small enough, one may consider the scattering as a point dipole field.¹⁵ One understands that the case in Fig. 5(a) is likely to be caused by a parallel (to the sample surface) exciting field at the very end of the tip, while the case in Fig. 5(c) corresponds to a vertical exciting field, and the case in Fig. 5(b) is a mix of the two polarizations. This understanding can be verified by the well known dipole field

$$\mathbf{E}(\mathbf{r}) = \frac{\omega^2 e^{i\omega/c_0 r}}{4\pi\epsilon_0 c_0^2} \left[\left(\frac{1}{r} + \frac{ic_0}{\omega r^2} - \frac{c_0^2}{\omega^2 r^3} \right) \vec{U} + \left(\frac{-1}{r} - \frac{3ic_0}{\omega r^2} + \frac{3c_0^2}{\omega^2 r^3} \right) \mathbf{nn} \right] \cdot \mathbf{p}, \quad (1)$$

where ω is the light angular frequency, c_0 the light speed in vacuum, \vec{U} a unit tensor, \mathbf{n} a unit vector in \mathbf{r} direction, and \mathbf{p} the dipole moment induced at the source point. However, in the above dipole field, there are two components: the homogeneous (propagating) field and the evanescent field. In the far field scattering, only propagating waves may be observed. In contrary, in near field optics, the desired part is the

evanescent field which dominates the field near the scatterer. Recently, one of us (M.X.) has realized an exact decomposition of the dipole field.¹⁶ The resulting evanescent field can be written as

$$\mathbf{E}_{\text{ev}}(\mathbf{r}) = \frac{\omega^2}{4\pi\epsilon_0 c_0^2} \left[\left(\frac{1}{2r} - \frac{c_0^2}{\omega^2 r^3} \right) \vec{U} + \left(\frac{1}{2r} + \frac{3c_0^2}{\omega^2 r^3} \right) \mathbf{nn} \right] \cdot \mathbf{p}. \quad (2)$$

The homogeneous field can be obtained either by a direct integration or by $\mathbf{E}_h = \mathbf{E} - \mathbf{E}_{\text{ev}}$ (see Ref. 16 for details). The importance of the above decomposed formalism to the tip testing is that the field distribution calculated with the above formulas are helpful to estimate the possible evanescent field distribution from the measured far field scattering pattern. It can be readily proven that the strongest evanescent field exists in the direction parallel to the exciting field. Therefore, the case in Fig. 5(a) means little evanescent field may interact with the sample features, while the case in Fig. 5(c) offers a strong concentration of the evanescent field towards the sample. In light of this, the tip in Fig. 5(c) may provide the best resolution in imaging. In practice, since the tip acts interactively with the sample features, the situation is more complicated than what we understand with simple dipole radiation and there is a need to have sufficient incident light. It has been found in our experiments that the tips between Figs. 5(b) and (c) would be the desired probes.¹² A detailed numerical study on the polarization effects in RSNOM is published elsewhere.¹⁷

In conclusion, in the present work we have developed a chemical etching process with a rotating chemical solution and multistep etching control, and we have made a new heating-pulling device to substitute the common laser-puller system. It is pointed out that as far as tips in RSNOM are concerned, chemical etching makes better tips. We have also

proposed a practical method to monitor the tip fabrication and to determine the usefulness of the tips by measuring the far field scattering pattern emitted from the very end of the fiber tip.

ACKNOWLEDGMENTS

The work was partially supported by CONACYT and DGAPA. Assistancess from the Supercomputer Center at UNAM is acknowledged.

¹Near Field Optics, edited by D. W. Pohl and D. Courjon (Kluwer, Dordrecht, 1993); for recent development in near field optics, see Ultramicroscopy **57**, 2/3 (1995).

²R. C. Reddick, R. J. Warmack, and T. L. Ferrell, Phys. Rev. B **39**, 767 (1989).

³E. Betzig and J. K. Trautman, Science **257**, 189 (1992).

⁴D. Courjon, J.-M. Vigourneux, M. Spajer, K. Sarayedine, and S. Leblanc, Appl. Opt. **29**, 3734 (1990).

⁵S. Bozhevolnyi, M. Xiao, and O. Keller, Appl. Opt. **33**, 876 (1994).

⁶G. Krausch, S. Wegscheider, A. Kirsch, H. Bielefeldt, J. C. Meiners, and J. Mlynek, Opt. Commun. **119**, 283 (1995).

⁷X. Zhu, G. S. Huang, H. T. Zhou, and Y. D. Dai, Solid State Commun. **98**, 661 (1996).

⁸T. Pangaribuan, K. Yamada, S. Jiang, H. Ohsawa, and M. Ohtsu, Jpn. J. Appl. Phys. **31**, L1302 (1992); T. Pangaribuan, S. Jiang, and M. Ohtsu, Scanning **16**, 362 (1994); M. Ohtsu, J. Lightwave Technol. **13**, 1200 (1995).

⁹G. A. Valaskovic, M. Holton, and G. H. Morrison, Appl. Opt. **34**, 1215 (1995), and the references therein.

¹⁰B. I. Yakobson, P. J. Moyer, and M. A. Paesler, J. Appl. Phys. **73**, 7984 (1993).

¹¹D. Zeisel, S. Nettesheim, B. Dutoit, and R. Zenobi, Appl. Phys. Lett. **68**, 2491 (1996).

¹²S. Bozhevolnyi, O. Keller, and M. Xiao, Appl. Opt. **32**, 4864 (1993).

¹³E. Betzig, J. K. Trautman, J. S. Weiner, T. D. Harris, and R. Wolfe, Appl. Opt. **31**, 4563 (1992).

¹⁴S. Bozhevolnyi, I. Smolyonimov, and O. Keller, Appl. Opt. **34**, 3793 (1995).

¹⁵M. Xiao, S. Bozhevolnyi, and O. Keller, Appl. Phys. A **62**, 115 (1996).

¹⁶M. Xiao, Opt. Commun. **132**, 403 (1996); J. Mod. Opt. **44**, 327 (1997).

¹⁷M. Xiao, Opt. Commun. **136**, 213 (1997).

Near-field optical microscopy with uncoated tips: Calibration, chemical contrast on organic crystals, and photolithography

G. Kaupp,^{a)} A. Herrmann, and M. Haak

University of Oldenburg, FB9 Organic Chemistry I, D-26111 Oldenburg, Germany

(Received 12 August 1996; accepted 14 April 1997)

Reflection-back-through-the-fiber scanning near-field optical microscopy (SNOM) measurements with uncoated tapered fiber tips, that remain cold, were performed under shear-force distance control for opaque and transparent rough samples. The feedback dynamics of the commercial (DME) RASTERSCOPE SNOM were tested with a 3000/500 nm gold grating at very high scan rates. Slow rates give rise to thermal broadening due to heating of the gold surface by partial light absorption. Enhancement of the light intensity at the base by reflection from steep walls on four sides is reported, the steepness contrast being less important. Doped gallium arsenide substrates give materials contrast. Extremely rough porous silicon has been probed. Organic chemical contrast is seen for the first time with SNOM on partly oxidized anthracene at large dislocations. Anthracene, anthraquinone, and their charge-transfer complex are clearly discriminated by different reflectivities. Irradiated α -cinnamic acid exhibits a chemically uniform surface and, thus, no SNOM contrast despite the anisotropic μm high features that are observed. Photolithography on light-absorbing 2,5-dibenzylidene-cyclopentanone crystals with the tip illuminated at an ≈ 10 nm distance provides an easy, inexpensive, and rapid performance test of single tips. Very long local irradiation creates features that widen and reach heights of 300 nm. © 1997 American Vacuum Society.

[S0734-211X(97)14504-8]

I. INTRODUCTION

Scanning near-field optical microscopy (SNOM) including light tunneling¹ was performed with eight basically different configurations.² The three most popular devices require very blunt metal coated tapered fiber tips that become very hot even at low light intensities and eventually melt at higher intensities.^{3,4} Such behavior is unsuitable for organic and biological samples that do not allow for temperature gradients of 100 K over distances of 10 nm or less and exhibit roughness. Therefore, cold, i.e., uncoated, tips are essential for organic samples. The best choice is the reflection-back-through-the-fiber mode⁵ if the feedback dynamics under shear-force control⁶ are high. Resolution below the diffraction limit is high even in very rough organic samples which cannot be probed by the more recent Lecher waveguide tetraprobe.⁷ Chemical and physical-state contrast may be obtained on natural samples if the reflectivity is discriminated.

II. EXPERIMENT

A RASTERSCOPE SNOM from DME (Copenhagen) was used with single mode uncoated tapered fiber tips (radius at the end 10–15 nm) that do not maintain polarization and that were pulled in a laser based P-2000 micropipette puller (Sutter Instruments, Novato, CA). Q factors were ≈ 100 and dither amplitudes were 10–30 nm at frequencies of 140–200 kHz near the mechanical resonance of the tips. Damping force settings were chosen as low as possible (0.012–0.017 nN), and loop gains of 0.125–0.177 were applied. The light intensity from an Ar⁺ laser (model 161 LGS from Laser Graphics) at 488 nm was measured with a Spec-

tra Physics powermeter model 404. The intensity that arrived at the end of the tip was 60–80 μW , the one coming back after passing the crossed polarizers to eliminate all unwanted reflections of primary light from lenses, interfaces, etc. was up to 1 μW . The scan rate was optimized between 10 and 120 $\mu\text{m/s}$ at a resolution of 512×512 points. The feedback dynamics were checked with a porous silicon sample that was covered with 50 nm of silica (nominal depth 2 μm , diameter 1.5 μm) and that gave only tip imaging when scanned with pyramidal Si_3N_4 tips. However, the fiber tip under shear-force control went down to more than a 1 μm depth between symmetrically traced hills at a 1.5 μm peak distance (50 of them in an area of 121 μm^2) at a scan rate of 40 $\mu\text{m/s}$ without significant overshooting. The measured roughnesses R_{ms} and maximum heights were 215 nm and 1.4 μm , respectively. The photomultiplier was from Hamamatsu, model H5783-01. Calibration for the width and height was done with an interferometric fabricated three-dimensional (3D) gold double grating (200 nm gold layer on silicon) type CAL 3000/500-A from IMS, Roskilde. The raw data were processed and imaged with Digital Instruments NanoScope III software.⁸

Single crystals of anthracene (1) were obtained from ethanol in the form of scales [$P2_1/a$, mainface (001)], the thickness of which was microscopically determined to be 8–10 μm . Monoclinic prisms of anthracene (1) were obtained from 1,2-dichloroethane [$P2_1/a$, mainfaces {110}]. The crystals were stored in air in closed bottles for 2 and 1 years, respectively (the prisms also for 2 months for comparison) prior to the SNOM measurements. The only product, anthraquinone (2), is well known and was identified by high-performance liquid chromatography (HPLC), and ultraviolet (UV), Fourier transform infrared (FTIR), ¹H-nuclear

^{a)}Electronic mail: kaupp@kaupp.chemie.uni-oldenburg.de

magnetic resonance (NMR) spectroscopy after chromatography of 1 g samples.

The charge-transfer complex (formula 3) was obtained by briefly melting a 1:1 mixture of compounds 1 and 2. After cooling, a homogeneous brown crystalline mass was obtained. The color disappeared upon solubilization, however solid state diffuse reflection spectra (Perkin-Elmer 551S UV/vis spectrometer with a reflection sphere and MgO as the white standard) indicated strong absorption at 488 nm for complex 3, whereas crystals of compounds 1 and 2 or ground mixtures therefrom were transparent.

Irradiations of crystals of α -cinnamic acid (4) were performed with a water-cooled mercury high-pressure lamp through a Pyrex filter from a 5 cm distance in vacuum or at 365 nm and a band pass of 6.4 nm. In photolithography the vibrating tip was kept at a damping force setting of 0.14 nN under shear-force control and positioned by proper offsets within a previously chosen area, prior to illumination for the times given.

III. RESULTS AND DISCUSSION

A. Gold and gallium arsenide substrates

Gold gratings are frequently used for calibration purposes. Light from illuminating fiber tips will however, locally heat up gold surfaces that absorb strongly at 488 nm. The standard reflectance for massive gold and vacuum deposited gold is 36.5% (480 nm) and 25.3% (490 nm).⁹ The thermal expansion effects that ensue decrease the resolution of the features and have to be minimized by very rapid measurements, which, of course, require high feedback dynamics. Figures 1(a)–1(c) show a selection of shear-force images of a 3000/500 nm gold double grating (prior to the final adjustment of the x, y, z piezos) with an illuminating tip at 100, 63, and 32 $\mu\text{m/s}$ scanning rates, respectively; the true shape of the features is shown in the perspective image of Fig. 1(d). Both types of images were taken for all rates with the same tip at the same settings with and without light in order to show that sharp features are obtained at high and at low rates in the absence of thermal effects. It can be seen that the very fast recordings do not exhibit overshoots. While the features in Fig. 1(a) approach, but do not fully reach, the sharpness of those measured with a nonilluminated tip [Fig. 1(d)], it is clearly seen in Figs. 1(b) and 1(c) that the shape of the features becomes more and more fuzzy the slower the scan rate is. Thus, even at rather dense features of 500 nm in height the very high scan rates are essential on gold surfaces if the quality of the SNOM image has to equal that of nonilluminated shear-force traces. Such heating problems do not occur, of course, with transparent or weakly absorbing samples.

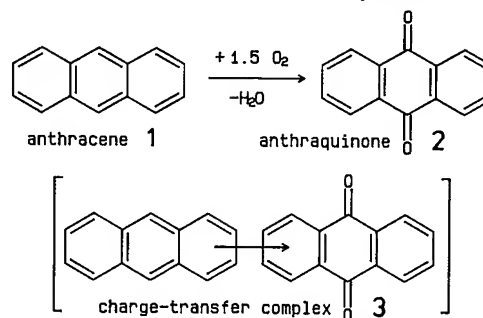
The gold double grating, which is scanned down to the flat base by our setup, is a valuable tool for calibration and for assessing topographic artifacts in reflection-back-through-the-fiber SNOM with slim uncoated tips. Figures 1(a) and 1(e) indicate attainment of the constant distance mode over all the grating. Thus, no SNOM contrast should occur on the rough gold surface. Figures 1(e) and 1(f) (thermal loss of resolution at the rate of only 70 $\mu\text{m/s}$) show,

however, that increased light pickup occurs if the tip is down above the flat regions between four adjacent steeply sloping gold planes that reflect light to it. Thus an increase of light intensity is observed due to far-field reflection from the 55° steep sides¹⁰ into the center. A steepness artifact will decrease the reflection into the fiber while scanning the slopes at constant distance because of sidewise reflection. But such artifacts are minor, with the bright SNOM features in Fig. 1(f) being only about five times as high as the peaks that are hard to see that correspond to reflections from the tops of the grating features in the low intensity regions. Important is the fact that in true constant distance mode with very slim tips, no topographic signals occur at 55° slopes or in features 500 nm in height. The far-field reflection from regularly fabricated gold surfaces is a unique feature.

In a similar way, a highly reflecting gallium arsenide substrate with dots of indium phosphide on it was very easily scanned with the RASTERSCOPE SNOM. Fewer reflecting and more reflecting spots side-by-side were detected indicating materials' contrast and its as yet unclear local chemistry (e.g., several phases, corrosion, etc.).¹¹ The relative reflectivities of materials will likely become available as soon as the technique matures.

B. Partially autoxidized anthracene

It appears to be advantageous to use the versatile SNOM setup for probing transparent organic crystals despite their low reflectivity. In the absence of thermal effects reflectivity differences of unlike molecules should give rise to chemical contrast. This has been achieved with crystalline anthracene,



which autoxidizes to give polar anthraquinone when large prisms from 1,2-dichloroethane (2 mm long and 0.5 mm thick) are exposed to oxygen, whereas very thin scales (8–10 μm) from ethanol (both have the space group $P2_1/a$) do not autoxidize.

Product 2 was analytically identified and further products were not found; however, we have synthesized the brown 1:1 charge-transfer complex (formula 3) that may be an intermediate and that absorbs the 488 nm laser light. Oxygen pre-treated scales (2 years) or prisms (1 year) of anthracene do not give SNOM contrast on their (001) and (110) faces. Thus, the flat surfaces with corrugations in the 10–30 nm range exhibit chemical uniformity and were not locally autoxidized. The reason for the nonreactivity towards oxygen on (001) and (110) is apparently the molecular packing of molecules 1. On (001) all molecules 1 stand steeply (67°) on their short edge.¹² Thus both central carbon atoms (the reac-

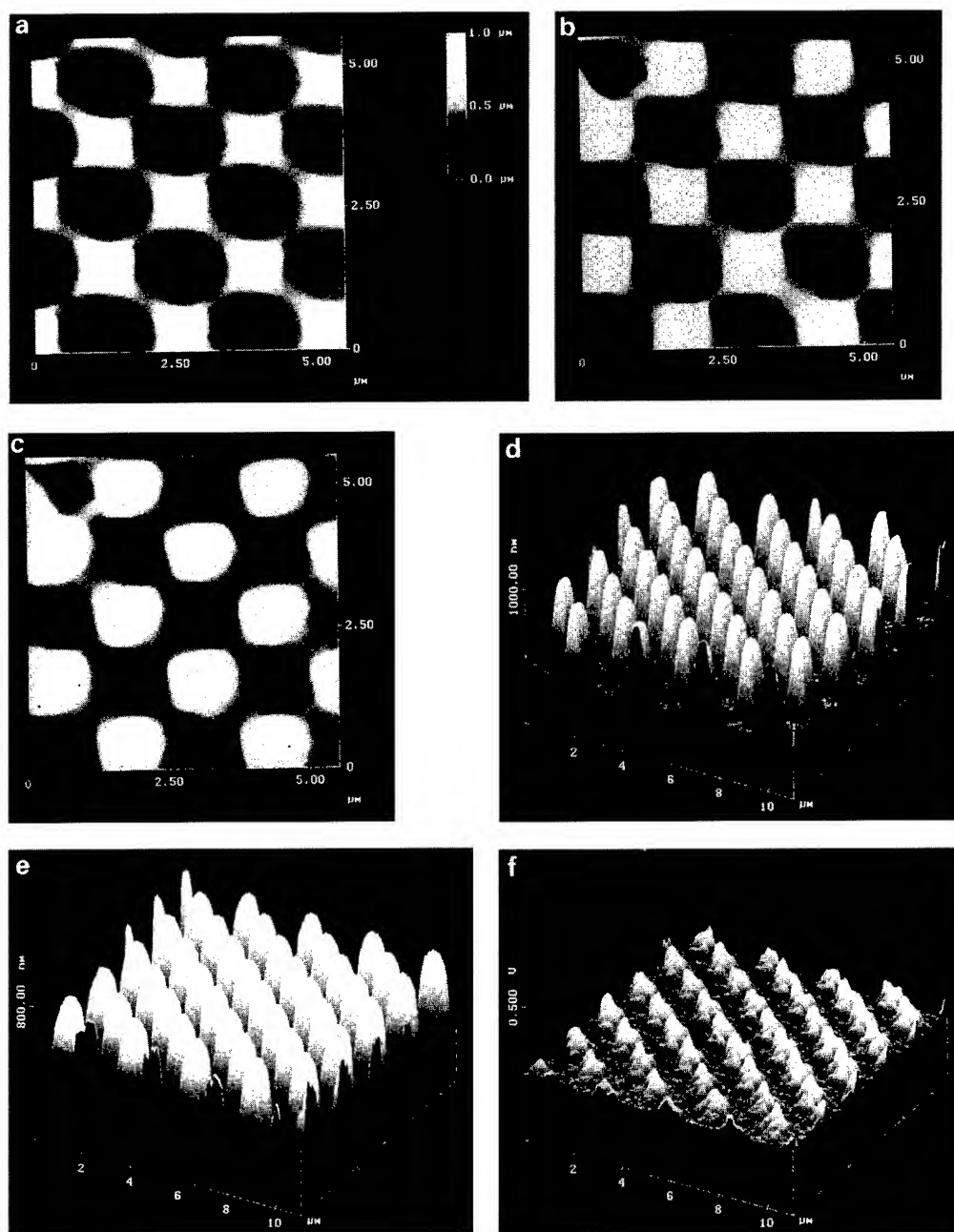


FIG. 1. 3000/500 nm gold double grating; shear-force topography with illuminating tip (488 nm, 60 μ W) at (a) 100, (b) 63, and (c) 32 μ m/s scan rates, showing the decrease in resolution due to a thermal effect; (d) shape of the undisturbed features measured with a dark shear-force tip at 40 μ m/s; (e) and (f) simultaneous shear-force AFM/SNOM images at 488 nm and a 70 μ m/s scan rate showing thermal broadening of the features and increased reflection yield when the tip is down between the features. All measurements were performed with the same tip.

tion centers) are completely hidden. On (110) molecules 1 stand on their long edge and are tilted at 76° and 23° ,^{12(a)} with the steep molecules shielding the ones lying flat. At best, one of the reaction centers is available from the molecules in the topmost layer and the multicenter reaction via the labile anthracene-9,10-endoperoxide also cannot begin. For geometric reasons, both reaction centers of anthracene will be available at faces that are skewed with respect to their natural {110} faces. That situation occurs at large dislocations, which are rather frequent on (110) of anthracene. One of them (1 year exposure to air) is shown in Fig. 2(a). The

height of the dislocation is 400 nm at a slope of 20° – 25° . Shear-force atomic force microscopy (AFM) reports only topography but cannot discriminate chemical identities, therefore the features at the base of the slopes of Fig. 2(a) do not appear unusual in AFM. It is only SNOM [Fig. 2(b)] that can determine chemical differences due to different efficiencies of reflectivity. The slightly corrugated planes (both levels) exhibit intermediate reflectivity. The features at the lower parts of the slopes and halfway up give enhanced reflectivity. Decreased reflectivity occurs at the upper parts of the slopes. So we have three different situations at the sites of chemical

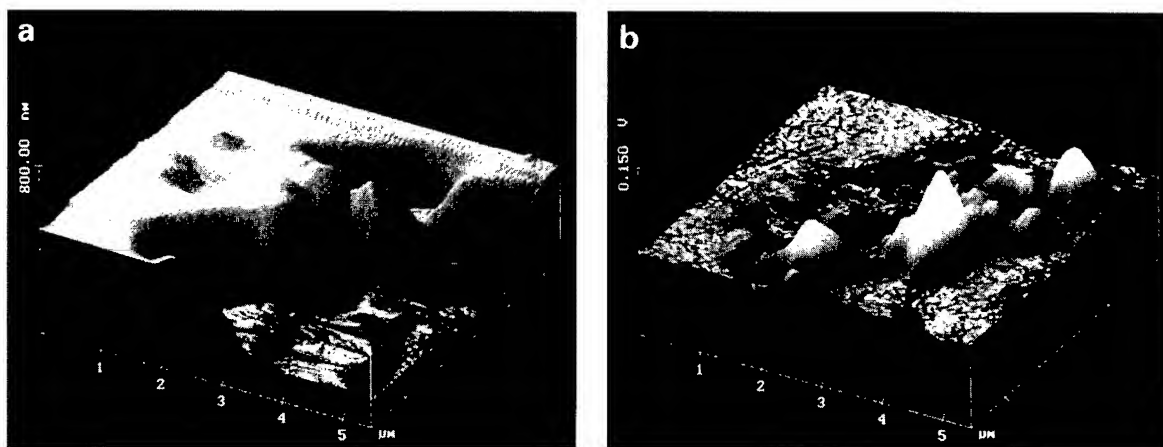


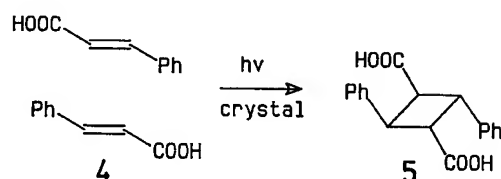
FIG. 2. (a) Simultaneous shear-force AFM and (b) 488 nm SNOM images at a large dislocation on a (110) face of 1 ($P2_1/a$) showing topography and chemical contrast, respectively; the scan rate was 50 $\mu\text{m/s}$.

reaction and correlation with the crystal structure is perfect. Assignment to the analytically identified species is straightforward. Unpolar anthracene (1) exhibits intermediate near-field reflectance. Polar anthraquinone (2) is more reflective in the near field. The light-consuming brown charge-transfer complex (formula 3) reflects into the tip the least. We are able to discriminate the sites of nonoxidation, full oxidation, and half-oxidation with submicroscopic resolution, and what we see is true chemical contrast. Similar behavior was found in prisms of anthracene that were exposed to air for 2 months at large (diameter 3 or 5.5 μm , depth 400 nm) or smaller craters in their (110) face. In those cases, the decreased near-field reflectance due to the sites of the charge-transfer complex (3) was more pronounced since the local oxidation was less complete. The bright SNOM features are very wide, however, and we can differentiate details of 100 nm in size in the increased reflectivity areas of the large scans.

The question of near-field reflectance efficiency is now being studied by Keller.¹³ Almost nothing is known yet, except that it is not the bulk refractive index that is important (quantum electrodynamics leads to the consideration of spacelike and timelike couplings). Answers to these problems have to be determined both theoretically and by reliable empiric measurements. That also requires local near-field spectroscopy, which cannot presently be performed on very rough organic crystals, but will probably emerge from the now commercial Courjon SNOM technique.^{5,12(b)}

C. SNOM on photolyzed α -cinnamic acid

Feature-forming photolysis of α -cinnamic acid (4) gives rise to α -truxillic acid (5) in a nontopotactic reaction.^{14,15}

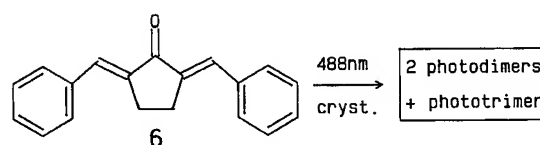


The reflectivities of conjugated compound 4 and nonconjugated compound 5 must be different. It was therefore of in-

terest to check if chemical contrast would be seen at and around μm high features that grow out of the bulk above the flat (010) surface, or if the chemical composition is uniform all over the surface. It can be seen from Fig. 3(a) that μm high features have built up anisotropically, with the slopes under the rims reaching 45° – 55° , as reported.^{14,15} The SNOM image, however, does not show significant contrast: the traces at the rims' location do not exceed the random roughness of the image. Again, we did not produce an optical image by scanning over topographical features which are more than 1 μm high. The same is true for 300-nm-high parallel ridges that were formed at a 365 nm tail irradiation and that do not exhibit SNOM contrast.¹⁵ These observations support the view that the photodimerization occurs everywhere in the surface region initially down through about 1300 molecular layers and that the primary feature-forming phase rebuilding is the result of pressure relief from the crystal bulk by long-range molecular movements along cleavage planes. Furthermore, it was shown that the very high features result from a secondary transformation. Nevertheless, the absence of a SNOM contrast indicates chemical uniformity all over the rough topography at least in the topmost layers of compounds 4 and 5. But chemical contrast is expected with all solid systems that choose an island mechanism.¹⁵

D. Photolithography with uncoated SNOM tips

As uncoated tapered tips remain cold, it is possible to produce photochemical feature-forming transformations with light-absorbing organic crystals^{15,16} rather than by ablating⁴ or melting of organic materials with hot coated tips. Thus, crystals of 2,5-dibenzylidene-cyclopentanone (6) photodimerize and -trimerize at 488 nm irradiation.



Such photochemistry with uncoated tips is certainly not a means for high-density data storage, because the illuminated



FIG. 3. (a) Simultaneous shear-force AFM and (b) 488 nm SNOM image of α -cinnamic acid on its (010) face after 15 min irradiation at $\lambda > 290$ nm showing up to $1.5 \mu\text{m}$ high features that grew anisotropically but do not exhibit significant SNOM contrast; the parallel fissures cut the c axis of the crystal at an angle of 40° .

spot is rather large. However, it does provide an easy and rapid means to judge the quality of single tips by the width of the features obtained (that is also important for metal-coated tips).

Furthermore, the directly obtained width of the illuminated area is important and was important for the interpretation of Fig. 1. Two hills have been created in Fig. 4 at different preselected offset positions. The large one grew to $1.4 \mu\text{m}$ diameter at 30 min irradiation by the nonscanning tip under shear-force control, while the smaller one acquired a 700 nm diameter after 5 min irradiation at a different preselected spot on a natural (001) face of compound 6. These features are indicative of a good tip. It is also shown in Fig. 4(b) that the large hills collected their material by forming a groove and by leveling down previously present features. Such lateral and vertical long-range molecular movements are common to organic solid state photochemistry. They choose the easiest ways that are offered by the crystal packing (e.g., along glide planes, etc.) and have been comprehensively elucidated by correlating AFM results with x-ray crystal structure data.^{15,16} Thus, for tip quality measurements a 2 min irradiation at $70 \mu\text{W}$ 488 nm output appears to be the best choice in our low cost, rapid test. Similar photolithography was reported recently using photodimerizing 9-chloroanthracene.^{15,16} It appears, however, that the easily available crystals of compound 6 are more versatile. The photochemical features may be also of interest for high-density data storage, but only if high-quality crystals are locally irradiated with low-aperture or apertureless⁷ tips. The features need not reach the enormous heights as in Fig. 4(b) and 3D data storage (location and height) should be feasible.

IV. CONCLUSIONS

SNOM with uncoated cold tips is a versatile tool that gives reproducible chemical contrast and chemically interpretable images at submicroscopic resolution beyond the diffraction limit. Very detailed conclusions can be drawn on rough surfaces just by recording relative reflectance. This technique will be quite important for use in organic and biological chemistry, in physics, and with inorganic materials. The light intensities that are collected at the recorder are probably high enough to use the reflection-back-through-the-

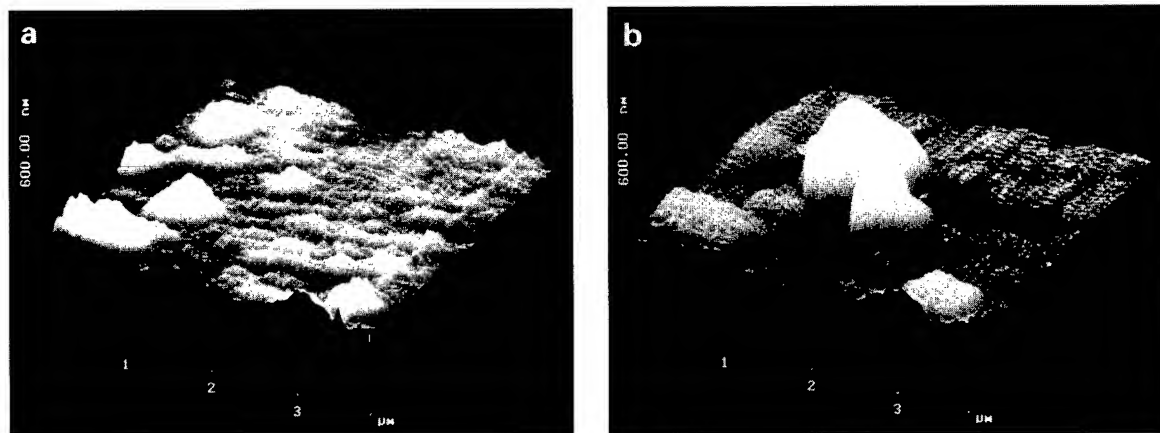


FIG. 4. (a) Shear-force AFM topography of a fresh (001) surface of compound 6; (b) the same surface after 30 min (large hill) and 5 min (small hill) irradiation at two preselected spots using an offset device under shear-force constant distance control showing an increase of the feature at long irradiation and the collection of material from the environment. The images were turned 90° with imaging software to show the groove that was also formed.

fiber SNOM in diffracting devices for local fluorescence and Raman spectroscopy of very high spatial resolution at fast recording while avoiding detrimental thermal effects and topology influences.^{12(b)} It is most important to have a technique that can be run without liquids, which, of course, change all kinds of solid organic surfaces by swelling, solvation, and dissolution at least on the molecular level.

ACKNOWLEDGMENTS

Thanks are due to Professor A. Smith, Leeds, for providing access to his SNOM equipment for one of the authors (A.H.), S. Madsen, Copenhagen, and the DME team for their initial help in running the RASTERSCOPE SNOM, and Professor H. Fuchs, Münster, for making his micropipette puller accessible. Dr. M. Schöning, Jülich, is thanked for the sample of porous silicon and the Deutsche Forschungsgemeinschaft for a travel grant.

¹E. A. Ash and G. Nicholls, *Nature (London)* **237**, 510 (1972); R. Young, J. Ward, and F. Scire, *Rev. Sci. Instrum.* **43**, 999 (1972); W. D. Pohl, W. Denk, and M. Lanz, *Appl. Phys. Lett.* **44**, 651 (1984); G. Massey, *Appl. Opt.* **23**, 658 (1984); U. C. Fischer, *J. Vac. Sci. Technol. B* **3**, 386 (1985).

²G. Kaupp and A. Herrmann, *J. Phys. Org. Chem.* (in press).

³A. H. LaRosa, B. I. Yakobson, and H. D. Hallen, *Appl. Phys. Lett.* **67**, 2597 (1995).

⁴D. Zeisel, S. Nettesheim, B. Dutoit, and R. Zenobi, *Appl. Phys. Lett.* **68**, 2491 (1996).

⁵M. Spajer, D. Courjon, K. Sarayedine, A. Jalocha, and J.-M. Vigoureux, *J. Phys. (France) III* **1**, 1 (1991); H. Bielefeldt, I. Hörsch, G. Krausch, M. Lux-Steiner, J. Mlynek, and O. Marti, *Appl. Phys. A* **59**, 103 (1994); G. Krausch, S. Wegscheider, A. Kirsch, H. Bielefeldt, J. C. Meiners, and J. Mlynek, *Opt. Commun.* **119**, 283 (1995).

⁶H. Heinzelmann and D. W. Pohl, *Appl. Phys. A* **59**, 89 (1994); E. Betzig and J. K. Trautmann, *Science* **257**, 189 (1992).

⁷J. Koglin, U. C. Fischer, and H. Fuchs, *J. Biomed. Opt.* **1**, 75 (1996).

⁸G. Kaupp, *J. Vac. Sci. Technol. B* **12**, 1952 (1994).

⁹Landolt-Börnstein; thermal conductivity of Au in films ≈ 6 , in polycrystals ≈ 7 cal/m s K: P. G. Borden, *J. Electrochem. Soc.* **127**, 2454 (1980); heat capacity: 0.0315 cal/g K: *Gmelin Handbook: Au* (Springer, Berlin, 1954), Vol. 3, p. 486; light intensity: 60–80 μ W typically on 8×10^{-9} cm²: this work.

¹⁰The slopes to the north and to the south are artificially traced steeper due to dithering in north/south direction when the scan proceeds from west to east, as usual.

¹¹Sample from the Max Planck Institut FKF, Stuttgart, courtesy of M. Schüttler; G. Kaupp, A. Herrmann, W. Rühle, M. Schüttler, and R. Rettig (unpublished).

¹²(a) G. Kaupp and M. Plagmann, *J. Photochem. Photobiol. A* **80**, 399 (1994); (b) we are working on the development of local near-field spectroscopy (Raman and fluorescence) with a charge coupled device (CCD) detection unit for the internal reflection SNOM to obtain more precise identification of the chemical species now identified by relative near-field reflectance; presently available systems are not suitable for very rough organic crystals.

¹³O. Keller, Fourth International Conference Near-Field Optics, Jerusalem, 9–13 February 1997, Book of Abstracts, NFO-4, Dan Knessim Ltd., Ramat-Gan; T. Andersen and O. Keller, in *Notions and Perspectives of Nonlinear Optics*, edited by O. Keller (World Scientific, Singapore, 1997).

¹⁴G. Kaupp, *Angew. Chem. Int. Ed. Engl.* **31**, 592 (1992); *Chemie Z.* **31**, 129 (1997).

¹⁵G. Kaupp, in *Comprehensive Supramolecular Chemistry*, edited by J. E. D. Davies (Elsevier, Oxford, 1996), Vol. 8 (6), p. 234; *Adv. Photochem.* **19**, 119 (1995).

¹⁶G. Kaupp, J. Schmeyers, M. Haak, T. Marquardt, and A. Herrmann, *Mol. Cryst. Liq. Cryst.* **276**, 315 (1996).

Preparation of probe tips with well-defined spherical apices for quantitative scanning force spectroscopy

U. D. Schwarz,^{a)} O. Zwörner, P. Köster, and R. Wiesendanger

Institute of Applied Physics and Microstructure Research Center, University of Hamburg, Jungiusstrasse 11, D-20355 Hamburg, Germany

(Received 8 September 1996; accepted for publication 5 March 1997)

A method for the preparation of scanning force microscopy (SFM) tips with spherically shaped tip apices and known tip radii by exposing commercially available silicon cantilevers to the electron beam of a transmission electron microscope is presented. The spherical shape of the tip apices was achieved by growth of a contamination layer at the end of the tip using the electron-beam deposition process. Well-defined radii between 7 and 120 nm could be produced. The importance of such tips for quantitative SFM measurements is discussed. Topographical measurements on a special test sample are shown as well as measurements of the frictional force as a function of the loading force as an example for quantitative spectroscopical measurements. © 1997 American Vacuum Society. [S0734-211X(97)02004-0]

I. INTRODUCTION

The principle of scanning force microscopy (SFM), originally developed by Binnig *et al.*¹ in 1986, relies on the accurate measurement of the forces interacting between a sharp, pointlike tip and a suitable sample. These forces can be of quite different nature such as attractive van der Waals forces, ion-ion repulsion, electrical, magnetical, capillary forces, etc.² In the years after the invention of SFM, the method was predominantly applied as a tool for spatially resolved *mapping* of forces. The main requirements for tips used in SFM experiments were stability during scanning (against wear and flexing) and a sharp tip apex (i.e., a small radius of curvature) in order to obtain a high lateral resolution and to prevent the occurrence of tip artefacts (see, e.g., Refs. 3–7).

In recent years, much effort was made towards exact quantitative measurements of the force acting between tip and sample (cf., e.g., Refs. 8–12). The most frequently used model for the comparison of such experimental quantitative force microscopy results with theoretical calculations is one of a perfect sphere with radius R interacting with an atomically flat surface; examples and illustrations of this issue can be found, among many others, in Refs. 2, 8, 13, and 14.

However, SFM tips, which are commonly used today, often show a significant deviation from the assumed sphere-on-flat model. It is obvious that a slightly truncated pyramidal tip will have a completely different contact area-load dependence than it will be if it is obtained from the sphere-on-flat model.^{15,16} This makes it difficult to obtain detailed quantitative information from SFM experiments on forces which depend on the actual contact area of tip and sample, as is the case, e.g., for adhesive, capillary, elastic, or frictional forces. Similar arguments also complicate the analysis of long-range forces such as van der Waals, electrostatic, or magnetic forces, which are mostly measured using noncontact modes. Therefore, many results from SFM measurements have so far

been restricted to qualitative imaging, keeping some parameters constant rather than reaching the goal of an absolute determination of the measured forces with a subsequent analysis in terms of a quantitative theory.¹⁷

The consequence of the issues discussed above is that for many quantitative SFM investigations, the use of tips with a known shape of the tip apices is mandatory. Since the tip/sample interaction area in force microscopy is often very small (contact areas below 1 nm² are possible even in contact mode measurements¹³), the tip shape should be exactly defined down to the nanometer scale. For the sake of easy comparison with theoretical calculations, a spherical tip apex is preferable. In this article, we will describe the preparation of SFM tips with a well-defined spherical shape of the tip apices and known tip radii which come quite close to the requirements mentioned above.

II. TIP PREPARATION

The preparation of tips with a well-defined spherical shape of the tip apices was realized by the contamination of doped single-crystalline silicon tips¹⁸ with amorphous carbon in a transmission electron microscope (TEM)¹⁹ by means of electron-beam deposition. The use of electron-beam deposition for the preparation of SFM tips is well known and many articles were already published describing various aspects of preparation and imaging properties of such tips (see, e.g., Refs. 20–24). These tips were always needlelike grown tips. The method presented in this article, however, includes the coating of a tip with a layer of amorphous carbon, as will be described.

To carry out the coating, the tips were positioned edge-on in the electron beam so that the complete cantilever was visible at low magnification in side view. This allows for the accurate measurement of the cantilever thickness, the cantilever length, and the tip length. These dimensions are, besides the width of the cantilever, important for a correct determination of the vertical and torsional spring constants of the cantilever.¹² Then, the electron beam was focused on the

^{a)}Electronic mail: schwarzu@physnet.uni-hamburg.de

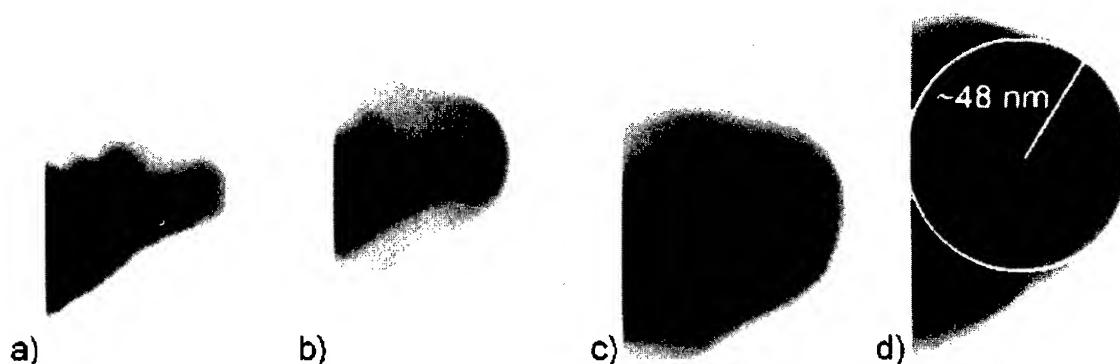


FIG. 1. Growth of a contamination layer on the apex of a commercially available SEM tip etched from doped single-crystalline silicon during the examination of the tip in the TEM. The micrograph presented in (a) was taken after 10 min exposure of the tip to the electron beam, (b) after 17 min (c) after 24 min, and (d) after 34 min. The radius of the tip at the beginning of the coating was estimated to ≈ 8 nm, the final radius was determined to (48 ± 4) nm.

very tip end at typical magnifications of about $430\,000\times$. The typical parameters of the microscope during imaging and coating were 60 kV acceleration voltage, $10\text{--}20\ \mu\text{A}$ emission current of the filament, $2\text{--}4 \times 10^{-4}$ Pa background pressure, and a spot size of $\approx 1\ \mu\text{m}$ in diameter.

During the imaging of the tip apex, the growth of a contamination layer can be observed (Fig. 1). Molecules from the residual gas are ionized in the electron beam and accelerated towards the tip end.^{20,22,23} There, the molecules spread out evenly due to their charge, forming a tip apex which comes very close to the shape of a perfect sphere. The final radius of the tip could be varied by the variation of total exposure time. Figure 1 shows the growth of a contamination layer after (a) 10 min, (b) 17 min, (c) 24 min, and (d) 34 min. Immediately after taking the electron micrograph shown in Fig. 1(d) the tip was removed from the electron beam to prevent further change of the shape of the tip apex. The tip radius can then accurately be determined from the electron micrograph. In the present example, the radius was estimated to be about 8 nm at the beginning and (48 ± 4) nm at the final stage.

Using the method described above, SFM tips with spherical tip ends possessing tip radii of $7\text{--}120$ nm were successfully produced (cf. Fig. 2). This enables SFM investigations not only with tips of known geometry, but also with the tip radius as an additional parameter which can be varied reproducibly within a large range. However, the production of tips

with large tip radii requires long exposure times of more than 1 h. Not only does the electric potential at the tip end decrease for larger radii, but also more material has to be accumulated to form a layer with a bigger radius.

Additionally, for the fabrication of tips possessing tip apexes which come close to the shape of a perfect sphere, it turned out to be essential that the original tip was conducting and sharp (tip radius ≤ 10 nm). Therefore, only tips etched from doped single-crystalline silicon¹⁸ were used for coating. If the tip was already slightly truncated at the beginning, all efforts failed to produce tips with a satisfactory spherical shape of the tip end. Therefore, it was not possible to recover formerly spherical tips which got truncated during extensive scanning by a second exposition to the electron beam either.

III. SFM MEASUREMENTS

In order to check the quality of the produced tips, we routinely imaged a special test sample. Small amorphous FeOOH particles, produced in an emulsion applying the preparation scheme described in Ref. 25, were precipitated from solution on a mica substrate. Subsequent heating to $\approx 200^\circ\text{C}$ was necessary to decompose residuals of surfactants used during the production of the FeOOH particles in the emulsion. The resulting FeOOH particles, which are distributed all over the mica substrate, are of excellent spherical shape (except some small ones which can be slightly de-

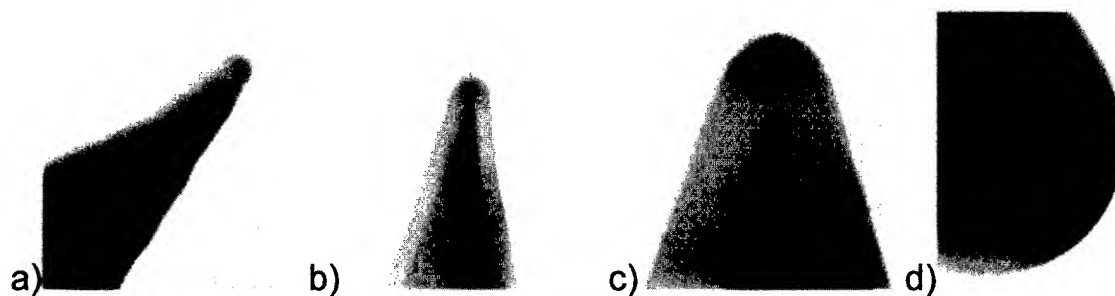


FIG. 2. TEM micrographs of some tip apices coated with a contamination layer produced by the described method. All tip apexes show nearly perfect spherical shape. The individual tip radii were determined to (a) (7 ± 2) nm, (b) (21 ± 5) nm, (c) (35 ± 5) nm, and (d) (112 ± 5) nm.



FIG. 3. SFM micrograph (image size: 900 nm×900 nm) of a test sample consisting of spherical FeOOH particles with diameters of 8–25 nm deposited on a mica substrate. The micrograph was acquired with a tip prepared as described in Sec. II; the tip radius was determined to (41 ± 5) nm. It can be seen that most particles are imaged with excellent circular shape, demonstrating the rotational symmetry of the tip apex.

formed) with diameters of about 8–25 nm. This test sample was well suited to check the shape of the tip apex and for a rapid estimation of the apex radius.²⁶

Figure 3 shows an SFM micrograph of the sample. The total image size is 900 nm×900 nm. It can be seen that most particles appear to be circularly shaped, which proves that the resulting tip apex has rotational symmetry despite the fact that the electron beam hits the tip from the side and the tip is not rotated during exposition. This was found for all tips prepared and can additionally be confirmed by taking cross sections in various directions through an individual particle.

A further test for the quality of the tip apex was performed by means of friction force spectroscopy. It can be shown that for a sphere with a small radius being in contact with a flat surface, the contact area-load dependence $A(F_n)$ follows $A \sim F_n^{2/3}$.^{13–15} Assuming a shear stress $S = F_f/A$, which does not depend on the mean contact pressure $p = F_n/A$, the frictional force F_f being observed when the tip is moved over the sample surface is also proportional to F_n to the power of 2/3:¹³

$$F_f \sim F_n^{2/3}. \quad (1)$$

We observed that this 2/3-power law is fulfilled very well on many materials (e.g., diamond, amorphous carbon, C₆₀ thin films, and germanium sulfide) as long as the tip apex shows no deviation from the spherical shape down to the nanometer scale in both electron micrographs and cross sections of the FeOOH test sample described above. An example is shown in Fig. 4. There, the (001) oriented surface of a single diamond grain of a diamond film epitaxially grown on silicon by chemical vapor deposition was investigated. However, as

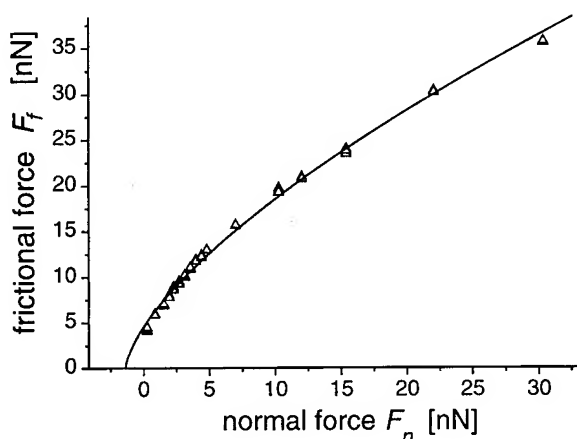


FIG. 4. Dependence of the frictional force F_f on the loading force F_n , measured on (001)-oriented diamond. Experimental and data analysis procedure are described in detail in Ref. 12. The $F_f(F_n)$ dependence follows in good approximation an $F_f \sim F_n^{2/3}$ law, as it is predicted for a small sphere over a flat surface in the case of a constant shear stress S (see Ref. 13).

soon as small truncations of the tip occur, significant deviations from the expected 2/3-power law are obtained. Additionally, all untreated tips which were used for friction force spectroscopy measurements also showed deviations from the 2/3-power law. This points to a tip-sample contact of the specially prepared tips which is close to the theoretical case of an ideal sphere over a flat surface and will be discussed in detail in a forthcoming publication.²⁷

A disadvantage for exact quantitative SFM investigations using these types of tips is the fact that the exact crystallographic structure and chemical composition of the tips are not known. The contamination layer is commonly assumed to be amorphous, which enables the formation of a spherical shape even on the nanometer scale. Its chemical composition depends on the composition of the residual gas in the microscope chamber. In Ref. 23, a contamination film of an electron microscope was investigated by electron probe microanalysis (EPMA), showing almost 100% carbon content and only small traces of oxygen. The presence of hydrogen is probable, but cannot be detected by EPMA. Consequently, the numerical values of Young's modulus E and Poisson's number ν , which are important for the determination of the tip/sample contact area using contact mechanical models such as the one described above, are unknown.

The wear resistance of the produced tips was found to be good: Tips could normally be used for some days even in contact mode before deviations from the spherical shape could be observed if the instrument was run in the regime of very low loading forces. This is in agreement with findings of other groups which also reported on a quite good wear resistance of tips which were produced by electron-beam deposition.^{22,24}

IV. CONCLUSIONS

To summarize, we presented a preparation procedure which allows the easy preparation of SFM tips with spherical

tip apexes by coating the tip end of commercially available silicon tips with a layer of amorphous carbon in a TEM by means of electron-beam deposition. Tip radii of 7–120 nm could be produced, which enables quantitative SFM measurements with tips of known geometry and the tip radius as an additional parameter. The application of such tips in topographical measurements is exemplified by scanning a special test sample of FeOOH particles which are deposited on a mica substrate; spectroscopical data of the frictional force F_f as a function of the loading force F_n are shown as an example for quantitative SFM measurements.

ACKNOWLEDGMENTS

The authors would like to thank M. Winzer and J. Müller for useful discussions, D. Lauter and I. Schneidereit for help with the TEM, W. Sager for the preparation of the FeOOH nanoparticles, and K. Schiffmann for the diamond sample. Financial support from the Deutsche Forschungsgemeinschaft (Grant No. WI 1277/2-2) is gratefully acknowledged.

¹G. Binnig, C. F. Quate, and Ch. Gerber, *Phys. Rev. Lett.* **56**, 930 (1986).

²For a review see, e.g., U. D. Schwarz, in *Handbook of Microscopy*, edited by S. Amelinckx, D. Van Dyck, J. Van Landuyt, and G. Van Tendeloo (VCH Verlagsgesellschaft, Weinheim, Germany, 1997), Methods II, pp. 827–844.

³M. Nonnenmacher, J. Greschner, O. Wolter, and R. Kassing, *J. Vac. Sci. Technol. B* **9**, 1358 (1991).

⁴D. Keller, D. Deputy, A. Alduino, and K. Luo, *Ultramicroscopy* **42–44**, 1481 (1992).

⁵J. E. Griffith, D. A. Grigg, M. J. Vasile, P. E. Russel, and E. A. Fitzgerald, *J. Vac. Sci. Technol. A* **10**, 674 (1992).

⁶U. D. Schwarz, H. Haefke, P. Reimann, and H.-J. Güntherodt, *J. Microsc.* **173**, 183 (1994).

⁷A. Khurshudov and K. Kato, *Ultramicroscopy* **60**, 11 (1995).

⁸J. L. Hutter and J. Bechhoefer, *Rev. Sci. Instrum.* **64**, 1868 (1993).

⁹D. R. Baselt and J. D. Baldeschwieler, *J. Appl. Phys.* **76**, 33 (1994).

¹⁰G. U. Lee, L. A. Chrisey, and R. J. Colton, *Sov. Science* **266**, 771 (1994).

¹¹E. Meyer, R. Lüthi, L. Howald, M. Bammerlin, M. Guggisberg, H.-J. Güntherodt, L. Scandella, J. Gobrecht, A. Schumacher, and R. Prins, in *The Physics of Sliding Friction*, NATO ASI Series E, edited by B. N. J. Persson and E. Tossati (Kluwer, Dordrecht, 1996), Vol. 311, p. 349.

¹²U. D. Schwarz, P. Köster, and R. Wiesendanger, *Rev. Sci. Instrum.* **67**, 2560 (1996).

¹³U. D. Schwarz, W. Allers, G. Gensterblum, and R. Wiesendanger, *Phys. Rev. B* **52**, 14976 (1995).

¹⁴K. L. Johnson, in *Micro/Nanotribology and Its Applications*, NATO ASI Series E, edited by B. Bhushan (Kluwer Academic, Dordrecht, 1997), Vol. 330, p. 151.

¹⁵K. L. Johnson, *Contact Mechanics* (Cambridge University Press, Cambridge, 1985).

¹⁶R. W. Carpick, N. Agrait, D. F. Ogletree, and M. Salmeron, *J. Vac. Sci. Technol. B* **14**, 1289 (1996).

¹⁷S. P. Jarvis, H. Yamada, S.-I. Yamamoto, and H. Tokumoto, *Rev. Sci. Instrum.* **67**, 2281 (1996).

¹⁸The cantilevers with integrated tips were provided by Nanosensors GmbH, D-71134 Aidlingen, Germany.

¹⁹Phillips, Model No. CM12.

²⁰T. Fujii, M. Suzuki, M. Miyashita, M. Yamaguchi, T. Onuki, H. Nakamura, T. Matsubara, H. Yamada, and K. Nakayama, *J. Vac. Sci. Technol. B* **9**, 666 (1991).

²¹K. L. Lee, D. W. Abraham, F. Secord, and L. Landstein, *J. Vac. Sci. Technol. B* **9**, 3562 (1991).

²²D. J. Keller and C. Chih-Chung, *Surf. Sci.* **268**, 333 (1992).

²³K. I. Schiffmann, *Nanotechnology* **4**, 163 (1993).

²⁴M. Wendel, H. Lorenz, and J. P. Kotthaus, *Appl. Phys. Lett.* **67**, 3732 (1995).

²⁵W. Sager, H.-F. Eicke, and W. Sun, *Colloids Surfaces A* **79**, 199 (1993).

²⁶C. Odin, J. P. Aimé, Z. El Kaakour, and T. Bouhacina, *Surf. Sci.* **317**, 321 (1994).

²⁷U. D. Schwarz, O. Zwörner, P. Köster, and R. Wiesendanger, *Phys. Rev. B* (in press).

Fabrication of a Si scanning probe microscopy tip with an ultrahigh vacuum-scanning tunneling microscope/atomic force microscope

Takahito Ono,^{a)} Hiroaki Saitoh, and Masayoshi Esashi
Faculty of Engineering, Tohoku University, Aza Aoba, Aramaki, Aoba-ku, Sendai 980-77, Japan

(Received 8 September 1996; accepted 10 March 1997)

Fabrication techniques for a sharp Si tip on a cantilever have been developed by using an ultrahigh vacuum-scanning tunneling microscope/atomic force microscope. The Si tip growth was performed by applying a voltage at a constant current between a Si substrate and the cantilever. It is considered that Si atoms were field evaporated on the substrate, and deposited onto the cantilever with this procedure. The sharp Si tip was successfully grown when a 3000 Å gold coated cantilever was used, and not grown when the gold film was absent. The cantilever that formed the tip may have applications in scanning probe microscopy. © 1997 American Vacuum Society.
[S0734-211X(97)02504-3]

I. INTRODUCTION

To obtain high resolution scanning probe microscopy (SPM) images, and to minimize a distortion of images depending on the tip shape, sharp tips are needed. Some techniques to fabricate sharp Si tips have been reported. The approaches to fabricate a sharp tip are divided into two methods, one is by etching and the other is by depositing. Batch fabrication of a Si structure based on an anisotropic etching, isotropic etching, dry etching, and oxidation for sharpening, is convenient to construct a sharp tip.¹⁻³ Techniques based on an electron beam deposition have also been reported.⁴ A high aspect ratio tip has been fabricated by using these techniques.

Recently, true atomic resolution has been demonstrated in noncontact atomic force microscopy (AFM).⁵ For atomic resolution in ultrahigh vacuum (UHV), we have developed a novel technique for fabricating a sharp tip in UHV. The tips were fabricated by depositing Si atoms by field evaporation with the UHV-scanning tunneling microscope (STM)/AFM. By using this technique, a sharp tip can be formed at the apex of the cantilever. Since this process is carried out in UHV, it can be used in the same chamber as the SPM probe, thus the probe doesn't have to be transported in the air for measurements. The Si tip is conducting, thus this probe can be applied to STM. In the case of noncontact measurement, the small curvature of the tip and high aspect ratio will enable high-resolution imaging.

II. EXPERIMENTS AND RESULTS

A. Ultrahigh-vacuum scanning tunneling microscope and atomic force microscope

All experiments of tip growth were performed with a JOEL JSTM/JAXFM4500XT unit. The substrate mounted on a holder is attached to a piezo actuator. This piezo actuator can be moved in the direction perpendicular to the substrate for up to 3 μm, and it is thermally insulated by ceramics, so

thermal drift is minimized. The growth of the tip can be monitored by a scanning electron microscope (SEM) combined with the UHV chamber.

B. Preparation of the Si substrate and the STM tip

A Si(100) substrate (*P* doped *n*-type 0.02 Ω cm) was heated by flowing direct current. Surface oxide was removed by heating the substrate at 1400 K for 30 s in UHV. The Au or W tips were electrochemically etched, and cleaned by heating in UHV. The temperature of the substrate was monitored by a pyrometer.

C. Growth of the Si nano-wire

A high aspect ratio structure could be formed on the tip by depositing atoms owing to a field evaporation of materials with the UHV-STM/AFM. First, experiments were performed by applying an appropriate voltage between the tip and substrate, and the substrate was heated, and then the tip-substrate distance was held to flow at constant tunnel current for a few minutes. Usually, the deposition occurs on the tip, and not on the substrate. Some combinations of substrate and tip material were tried as described below.

1. Gold tip and a Si(100) substrate

A Au tip and a Si(100) as a substrate were used in the first experiments. The substrate was heated to 700 °C. The applied sample voltage was 5 V and the tunneling current was 10 nA. The voltage was applied for a period of 15 min. Under the above conditions, the tip-sample distance was about 6 Å. The deposition on the Au tip could be monitored by the *in situ* SEM. The deposition shows a wirelike structure. The growth rate of the wire was not reproducible, however, the reason for these variation was unclear. The maximum growth rate was 6.0 nm/s with 150 nm diameter of the wire. The minimum diameter was below 150 nm with 3 μm length as shown in a transmission electron microscope (TEM) image in Fig. 1. It was revealed that the deposition consists of Si atoms by energy-dispersive x-ray analysis

^{a)}Electronic mail: tono@cc.mech.tohoku.ac.jp

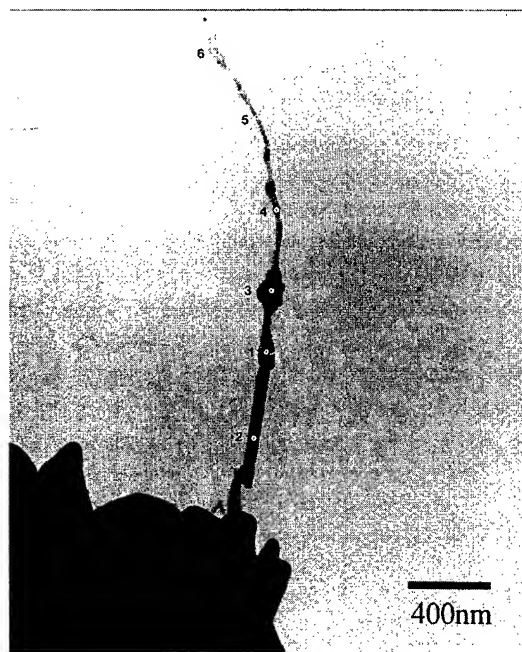


FIG. 1. TEM image of the wirelike structure grown on the gold tip. The deposit is longer than $3\ \mu\text{m}$, and a diameter is less than $150\ \text{nm}$. By EDX analysis, the deposit mainly consists of silicon. It is found to be crystallized by evacuation of the electron diffraction pattern.

(EDX) and TEM as mentioned later. When a Si(111) substrate was used as a substrate in the same experiment, deposition was also observed on the tip.

2. Tungsten STM tip and a Au thin film formed on a Si(100) substrate

Au sputtered Si(100) substrate was used. The Au thickness was $\sim 3000\ \text{\AA}$. Before forming the Au film, natural oxide was removed by ion gun sputtering. Using this substrate and the W tip, experiments were performed under similar conditions in which the sample bias and the tunnel current were $5\ \text{V}$ and $10\ \text{nA}$, respectively. The growth of deposition was observed on the W tip. Figure 2 shows the SEM image of the deposition which was grown for $15\ \text{min}$.

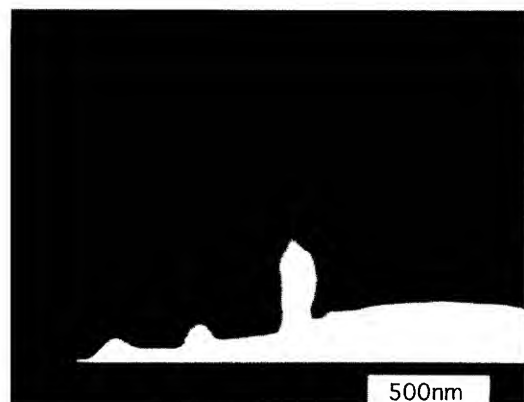


FIG. 2. SEM image of the deposit on the tungsten tip: The gold film formed on the Si substrate and tungsten tip were used for growth.

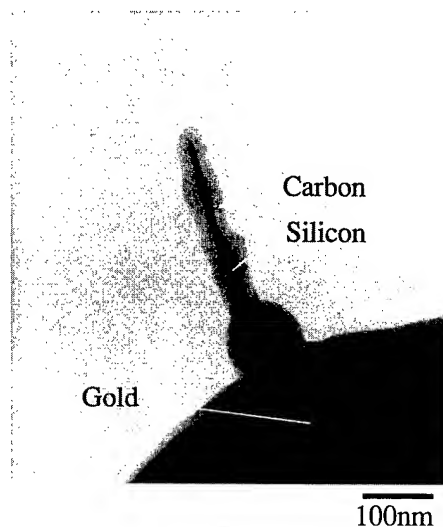


FIG. 3. TEM image of the deposit: It shows a solid cylindrical structure; the core and tube consist of silicon and carbon, respectively.

3. Tungsten STM tip and a clean Si substrate

When a W tip and a clean Si(100) substrate were used, deposition could not be observed by SEM. The experiment was performed under the same conditions as described in experiments (a) and (b). It is suggested that the Au atoms play important roles in the formation of the wirelike structure.

D. Evaluation of the Si nano-wire with transmission electron microscope

By observation with TEM (200 kV TEM HITACHI HF-2000), it was found that the maximum length of deposition was longer than $3\ \mu\text{m}$. In most cases, the diameter of the deposition becomes narrower as the growth continues. Deposition shown in Fig. 1 has been investigated by EDX. It was found that the deposition mainly consists of silicon. From electron diffraction analysis, diffraction patterns obtained at $200\ \text{keV}$ show diamond structure. The lattice constant is within 1% of that of bulk silicon.

Figure 3 shows the TEM image of deposition grown on the Au tip. This deposition shows a filled cylindrical structure which is composed of a core and a tube. EDX analysis revealed that the core and tube consist of silicon and carbon, respectively. It is not clear if the cylindrical structure was formed. There is a possibility that the structure was caused by contamination of the Au STM tip. By using a clean Au tip, such cylindrical structure was not observed. Nevertheless, this long wire is useful for observation of STEM images. We have obtained stable atomic images of Si(111) 7×7 using this wire as a STM tip.

E. Fabrication of SPM probe

Fabrication of the tip on the Si cantilever by field evaporation deposition was carried out in our UHV-STM/AFM.

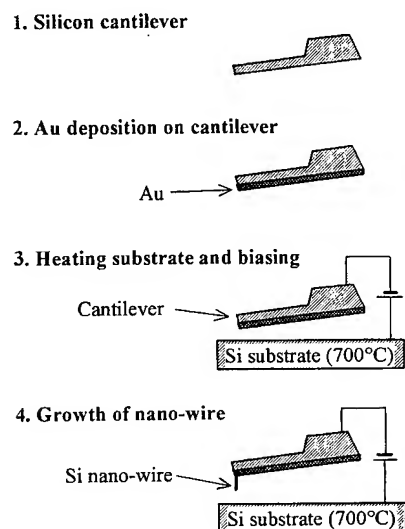


FIG. 4. Schematic diagram of fabrication of the tip: A gold film of 3000 Å thickness was formed on the cantilever, and the substrate was biased at 5 V and a tunnel current of 10 nA with a heating substrate at 700 °C.

Commercial Si cantilevers and fabricated ones were used. In both cases, the tips were successfully grown on the cantilevers by procedures described below.

The cantilever was fabricated by anisotropic etching of Si(100) wafer. After having been oxidized and prestructured by photolithography, the cantilevers were etched for cantilever thickness in tetramethyl ammonium hydroxide [TMAH, $(\text{CH}_3)_4\text{NOH}$] solution. The wafers were reoxidized and patterned with photolithography, and etched in TMAH solution to release the structures. The experiments were carried out using two methods. In method I, the cantilever was coated with gold while the substrate was not. In method II, the substrate was coated with gold while the cantilever was not. The experimental procedures of the two methods are described below.

Method I. The process flow chart for this method is shown in Fig. 4. Before Au deposition, natural oxide was removed from the Si cantilever by ion milling. 3000-Å-thick Au film was deposited on the surface of the Si cantilever. The tip was then transferred to the UHV-STM/AFM chamber with a probe holder for the tip fabrication. Bias voltage applied to the substrate was 5 V and a tunnel current was 10 nA. The substrate was heated to 700 °C for 10 min during the tip fabrication. In this procedure, the deposition was observed on the flat cantilever. The length, width, and thickness of the fabricated cantilever are 90, 20, and 2 μm, respectively. The spring constant is 7.13 N/m, and the resonant frequency is 299 kHz. The SEM images of the cantilever and the tip are shown in Figs. 5(a), 5(b), and 5(c). In these images, a tip with a height of 300 nm and a diameter less than 100 nm, is observed. In Fig. 6, another image of the fabricated tip is shown. This tip is grown on a commercially available cantilever. This picture suggests that the deposits were grown at protruding locations where the field evaporation occurs.

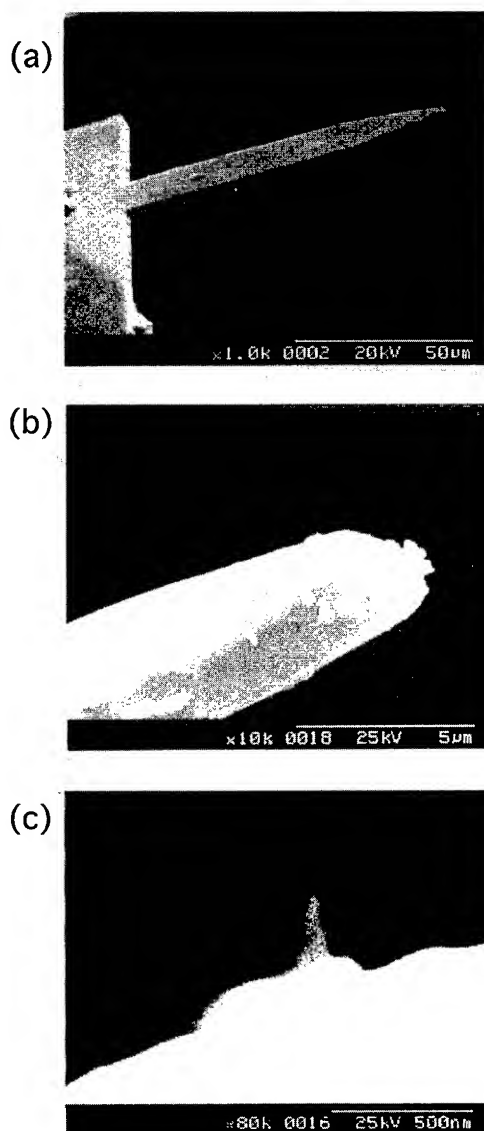


FIG. 5. (a) SEM image of cantilever which has a length of 90 μm, a width of 20 μm, and a thickness of 2 μm. Thin gold film was formed on the cantilever before tip growing. (b) Magnified image of the tip region of the cantilever. (c) The tip fabricated on the flat cantilever, which shows a height of 300 nm, and a diameter of less than 100 nm.

Method II. The procedure similar to method I is carried out with an Au deposited substrate. The cantilever was not coated with Au. It was observed that the deposition was grown. However, the deposition formed with method II was rather big or forming a lump that is grown at other locations. It is considered that the formation and condensation of silicide at 700 °C have caused such shapes. Another possibility is that the areas surrounding the cantilever may occasionally be touched since the cantilever is relatively flat. The deposition tip becomes sharper as the grown continues. Such a tip shape seems to be quite suitable for the SPM probe application. Comparing the results of the two experimental methods, method I seems to be appropriate for fabrication of the Si tip.

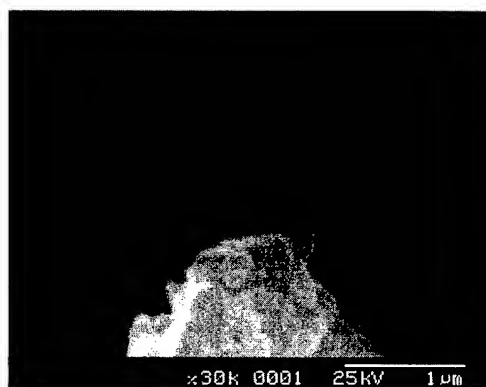


FIG. 6. The deposit was grown on the tip fabricated cantilever on which gold film was formed. The deposit is grown on the protruding point of the tip close to the Si substrate.

III. DISCUSSIONS

A. Electric field evaporation

Field evaporation is a phenomenon in which atoms are ionized and extracted from a surface by a high electric field. It is thought that there is some probability of deposition on a tip of the extracted atoms.⁶ Thus, the structure may be fabricated by depositing field evaporation. However, field evaporation is improbable, thus the deposition rate must be small. Field evaporation rate is expressed with an Arrhenius equation, $R = \nu \exp(-Q/kT)$, where ν is the vibration frequency, Q the activation energy, k the Boltzmann constant, and T the temperature. Thus it is found that the field evaporation rate should increase with decreasing activation energy and increasing temperature.

B. Si nano-wire growth mechanism

When voltage is applied between an Au tip and a Si substrate, because of field evaporation the atoms from the Au tip are transferred to the substrate while the atoms from the substrate are transferred to the tip. In this case, the Au atoms deposited on the substrate form Au silicide. Au silicide probably decreases activation energy, thus many Si atoms are deposited on the tip. In addition, if temperatures of the tip and substrate are different, the field evaporation rate be-

comes higher at the high temperature side than that at the low temperature side. For this reason, atoms will be deposited only at the low temperature side.

C. Applications as an SPM probe

For the methods described above, the tip can be grown on the cantilever adjacent to the substrate in a self-assembly manner. It is believed that the fabricated tip can be used as a high resolution SPM probe. Some potential advantages for using this probe are:

- (a) It can be used for an STM tip because the tip is conductive.
- (b) High aspect ratio tip can be obtained. It is expected to obtain a high resolution image by using the tip.
- (c) The tip can be fabricated at relatively low temperature in the UHV chamber which is used for SPM.
- (d) An extremely small cantilever will be fabricated. If the cantilever beam is small, the resonant frequency becomes higher, thus high speed image will be obtained.

IV. CONCLUSION

We developed novel technique to fabricate the Si tip by depositing with UHV-STM/AFM. The deposits of Si atoms having a wirelike structure have been formed on the Au tip by using the field evaporation method. It is found that the Au atoms contribute to the deposition of the Si atoms. By using this technique, the SPM probe can be fabricated by growth of the Si tip on the Au-coated cantilever.

ACKNOWLEDGMENTS

The authors would like to thank K. Takeo, M. Hiroaki, and T. Yaguchi (Hitachi Instruments Engineering Co.) for TEM analysis.

¹S. Akamine, R. C. Barrett, and C. F. Quate, *Appl. Phys. Lett.* **57**, 316 (1990).

²T. R. Albrecht, A. Akamine, T. E. Carver, and C. F. Quate, *J. Vac. Sci. Technol. A* **8**, 3386 (1990).

³A. Boisen, O. Hansen, and S. Bouwstra, *J. Micromech. Microeng.* **6**, 58 (1996).

⁴D. J. Keller and C. C. Chung, *Surf. Sci.* **268**, 333 (1991).

⁵F. J. Giessible, *Science* **267**, 68 (1995).

⁶D. Huan, H. Uchida, and M. Aono, *J. Vac. Sci. Technol. B* **12**, 2429 (1994).

Conductive supertips for scanning probe applications*

C. Schössler and J. Urban

Institut für Angewandte Physik der Technischen Hochschule Darmstadt, D-64289 Darmstadt, Germany

H. W. P. Koops^{a)}

Deutsche Telekom AG, Technologiezentrum, D-64295 Darmstadt, Germany

(Received 8 September 1996; accepted 4 April 1997)

Materials machining with the scanning tunneling microscope suffers from poor linewidth compared to the atomic resolution power of the microscope. The trace of the emitted beam is widened due to electron or ion field emission from many tip locations having a low work function. To improve the result emission from a single site that delivers a beam in a confined emission angle is required. Such an emitter is obtained if a supertip is used. It consists of a blunt base tip and an attached supertip of a few nanometers in diameter and height. The supertip delivers the current from one point of field instability only. The attached miniaturized tip generates the high field required for field emission. Electron-beam induced deposition from organometallic gold compounds and a heated substrate is used to build the attached nanocrystalline supertip. Confinement of the emission angle of the emitted beam is confirmed by field emission microscope investigations. An angular confinement of $\pm 7.2^\circ$ is obtained. Supertips deliver an emission of 0.2 mA/sr as measured, and have therefore at least a tenfold higher angular emission density than conventionally etched tips. Deposited supertips require no single crystalline base and can be placed on any base material. They successfully operate in a scanning tunneling microscope in air. © 1997 American Vacuum Society.

[S0734-211X(97)04904-4]

I. INTRODUCTION

Scanning tunneling microscopy (STM) applications for materials analysis and machining as well as for lithography employ field emission to use the highest brightness obtainable. All field emission tips suffer from low angular emission density, which is due to electron or ion emission from many emission sites on the tip operating in parallel. The current is emitted from areas of low work function and microprotrusions which generate local field enhancement. These sites are distributed over the sphere of the tip and produce the widespread angular emission. This results in poor resolution of STM lithography compared to the atomic resolution obtained in imaging. A much higher voltage than in the scanning probe techniques must be used to obtain a beam with field-electron emission or field ionization. For lithography a beam having a few eV of energy is required in order to introduce radiation damage in a resist film covering the sample. For a spherical tip especially both conditions result in a wider tip to sample separation and allow parallel emission of several points of the tip surface. This widens the profile of the emitted beam. A further spread of the beam occurs since the beam has to travel a few nm in air or in vacuum to reach the sample.¹ If the tip is formed while cutting a wire, a ridge that can operate as a tip is obtained; see Fig. 1(a). This results in astigmatic performance in lithography as is seen by the exposure of hydrogen activated silicon with a 4 eV beam from a Pt/Ir tip in air; see Fig. 1(b). To improve resolution of the STM lithography it is necessary to use sharper tips that produce confined beams.

This article describes the reproducible construction of conducting nanocrystalline supertips on top of conventionally etched base tips. The emissions characteristics are investigated and an application for STM is shown.

II. SUPERTIPS BY NANOLITHOGRAPHY WITH ELECTRON-BEAM INDUCED DEPOSITION

To build three-dimensional nanostructures electron-beam induced deposition is employed.² The electron beam of a scanning electron microscope with a field emission cathode is used to destroy organometallic molecules that are issued from a reservoir with a nozzle of low conductance to the sample and are adsorbed at its surface. The reservoir and nozzle are attached to the sample holder and both are inserted into the microscope via the specimen exchange airlock. By employing organometallic materials that contain gold, resistors and field emitter tips are deposited. High resolution and high aspect ratio structures are grown by this technique.³ The resistivity of gold containing resistors can be controlled over seven orders of magnitude by changing the beam current and the sample temperature. The metal content normally is in the range of 1–15 at. %, but can be increased using a higher substrate temperature than room temperature. The mechanism of conduction is electron hopping from crystallite to crystallite in the nanocrystalline material.⁴ The highest resolution obtainable is achieved using the electron beam of a scanning electron microscope equipped with a cold field emission cathode. Having computer control and image processing attached to the microscope three-dimensional nanostructures are generated using a unique macrocontrol of the beam position and exposure time for each pixel. Field emitters are deposited from dimethylgold-(trifluoro)-acetylacetonate [$\text{Me}_2\text{Au}(\text{tfac})$]. Earlier investiga-

*No proof corrections received from author prior to publication.

^{a)}Author to whom correspondence should be addressed.

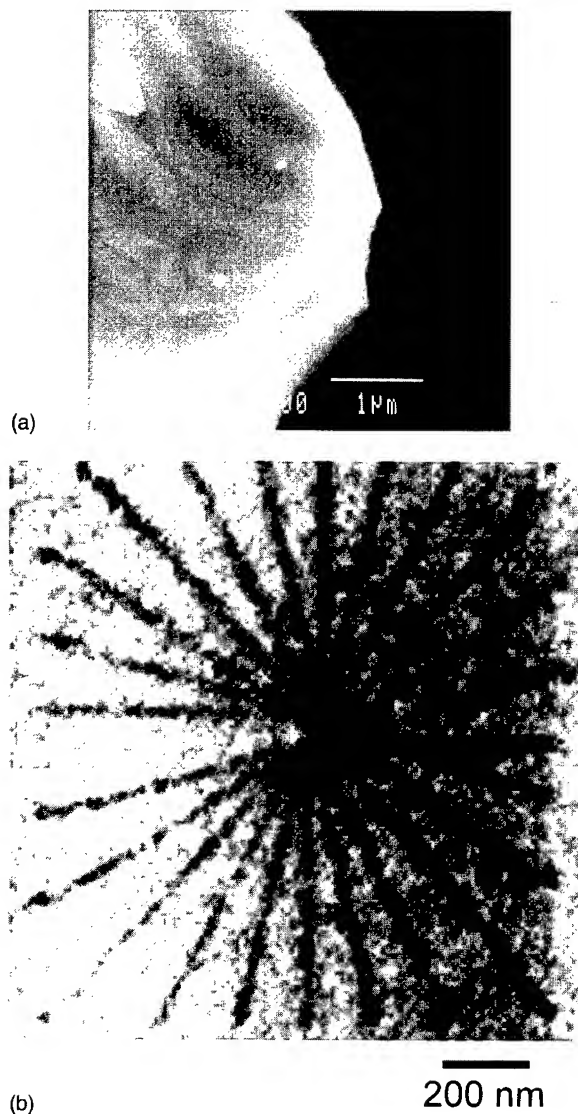


FIG. 1. (a) Scanning electron microscopy micrograph of the top part of a freshly cut Pt/Ir wire to be used for STM lithography in air. (b) Lithography with a cut Pt/Ir wire tip on hydrogen passivated silicon in air is possible using a positive tip voltage. The thickness of the oxide deposit is 1.2 nm.

tions of deposited tips in a transmission electron microscope revealed that single crystallites of nanometer size of the respective metal are embedded in an amorphous carbon containing matrix.⁴ For special deposition conditions the deposits form crystallites which protrude from the tip itself like teton tips. This leads to an increased field enhancement factor and lowers the extraction voltage for field-electron emission. A high emission current is accomplished due to the multiple emission centers available at the front of the tip.⁵ Deposited tips emit high field emission currents at low extraction voltages. Using a special arrangement with a only a 300 nm distance between the emitter and anode, earlier experiments showed field emission starting at a voltage as low as 10 V. At 22 V extraction voltage the emission current was in excess of 180 μ A for several minutes before the wire recrystallized and bent due to the high power load.⁴ The deposition of supertips at room temperature showed no re-

producible conductive connection to the substrate. This, however, was observed using platinum deposits to repair gated field emitters.⁶

A. Deposition of supertips on etched tungsten tips at elevated temperature

To investigate the emission image of the deposited tips, thin needles are produced from $\text{Me}_2\text{Au}(\text{tfac})$ on short etched tips made from tungsten wire. The tips are welded to a directly heated tungsten hairpin or are clamped into a carbon vise. A hairpin or a vise is used for resistive heating of the tip. The welded wires had mechanical imperfections that could not be corrected so they could not be used in the STM tip holders. Clamping the etched wire tips in a carbon vise rendered them easy to handle and produced good results. Before deposition, a 2000 °C thermal treatment of the tip is performed to remove the oxide from the tungsten tip. Since the flash treatment is performed in the presence of the organometallic molecular stream impinging on the tip, this is coated with a thin film of pyrolytic gold. This ensures good electrical contact between the deposited tip and the tungsten tip. Deposition was carried out at a sample temperature ranging from 20 to 100 °C; however, at the elevated temperature the growth rate of deposits is reduced. With this treatment at higher temperature electrical conducting deposits are routinely obtained. They were measured from the resistivity of the deposited wires. Energy-dispersive x-ray analysis of materials deposited at higher substrate temperatures revealed an increase in the gold content and a decrease in the carbon content. A comparative study of the deposited tips was performed with a high voltage and high resolution transmission electron microscope. The observations reveal that tips deposited at room temperature have a fine grain polymeric foot that is topped by a nanocrystalline tip with crystals of 4 nm

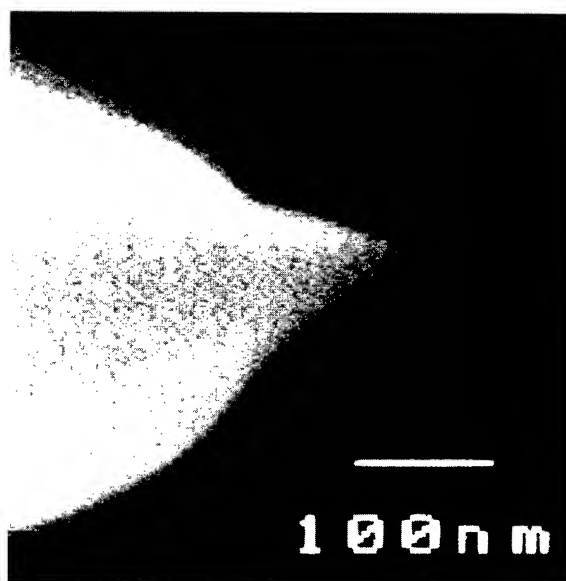


FIG. 2. Deposited supertip from $\text{Me}_2\text{-Au-tfac}$ on top of an etched tungsten wire. The tip radius is less than 15 nm.



FIG. 3. Emission pattern of an etched W (110) tip imaged in a field-emission microscope. The emission fills an angle of $\pm 30^\circ$. The inset shows an emission pattern of a gold supertip deposited at 25 kV and 100 pA at a sample temperature $> 100^\circ\text{C}$. The beam fills an angle of $\pm 7.2^\circ$.

diameter. Tips deposited at 80°C and higher temperature have a nanocrystalline appearance with crystals of 4 nm diameter all over the tip shank and up to the top of the tip. Gold-containing supertips with radii of curvature of some nm are deposited in this way. Figure 2 shows a supertip with a tip radius of 15 nm.

B. Field-electron emission characteristics

Conductive tips deposited as supertips on top of etched tungsten tips were analyzed in a field electron microscope. Emission patterns were observed operating the field-electron microscope in near ultrahigh vacuum. The emission images produced on a phosphor screen 25 mm from the tip are re-

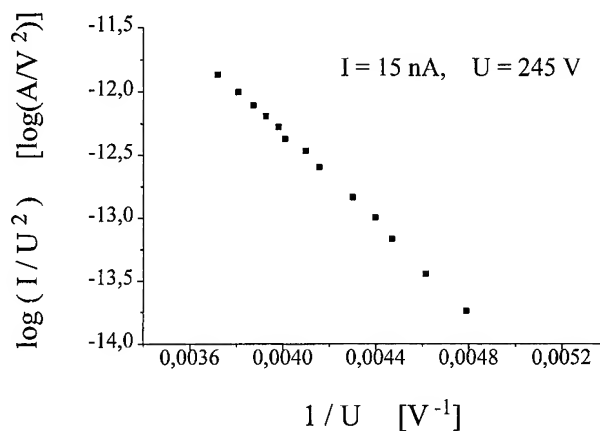


FIG. 5. Fowler-Nordheim graph of the emission current of a deposited gold supertip.

corded with a charge coupled device (CCD) camera. An aperture of 2 mm bore diameter is used as the extractor. The tip is centered to the bore and located at its front plane, that is, a ring-shaped electrode is operated at higher positive voltage than the screen anode placed at half the distance between the tip and the screen. This arrangement acts as an ion mirror by producing a positive saddle point for the potential between the cathode and the screen. If it keeps ions that are produced at the anode from running into the emitter tip.

Ordinary tungsten tips were imaged to test the arrangement. The experiment showed the well known emission image of (110) oriented tungsten tips at 2.5 kV extraction voltage, which is characterized by twofold symmetry. The diameter of the image was 40 mm corresponding to $\approx 30^\circ$ emission angle; see Fig. 3.

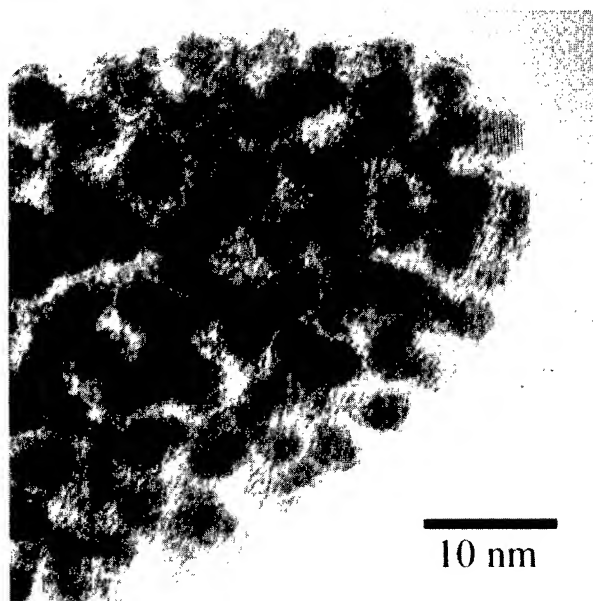


FIG. 4. The high resolution transmission electron micrograph of the tip of a supertip reveals the nanocrystallinity of the material. Perfect single crystals grow under the influence of substrate temperature and beam impact.

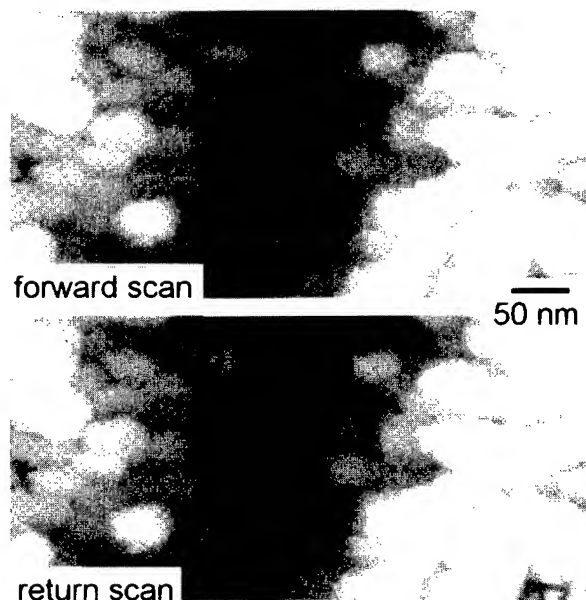


FIG. 6. STM image of a vacuum deposited gold surface taken in air at 500 mV using a gold supertip on a tungsten wire shows the stable action of the supertip.

C. Field-emission microscope characterization of deposited supertips

To characterize deposited supertips prior to STM use we have selected a field-emission microscope to investigate the electron emission properties. A system with an airlock for the fast exchange of tips was employed. Deposited supertips with nanocrystalline field emitters start field-electron emission at 800 V and reach a stable emission at voltages as low as 210 V by a forming process.⁷ Tips imaged in a SEM after field formation exhibit a deposit that is reduced in height and has a rounded apex radius. The lifetime of deposited field-emission tips was more than 2 h, which was the duration of the measurements. A stable field emitter with stable emission characteristics was obtained. The emission from one single site was regularly confined to an angle of $\pm 7.2^\circ$, as was observed on the screen; see inset of Fig. 3. This beam confinement to one emission site with a narrow angle is demonstrated for the first time for a deposited supertip. Similar confinement was reported earlier for field ion emission sources⁸⁻¹⁰ and for field electron emission.¹¹

The high resolution transmission electron micrograph of the tip of a supertip reveals the nanocrystallinity of the material, shown in Fig. 4. Perfect single crystals grow under the influence of substrate temperature and beam impact.

Emission currents ranged up to 10 μA from the confined emission site which corresponds to an angular emission of 0.20 mA/sr. This means that employing deposited supertips provides cold, single emission site field emitters with an angular emission density increased by a factor of 10 compared to conventional tips. The emission current was recorded and it generally followed Fowler–Nordheim characteristics; see Fig. 5. A measurement of the emission current of 1 μA versus time displays short-time variations of the order of 10% after forming the tip.

A sputtered gold surface was imaged with a STM in air at 500 mV using a deposited supertip. Figure 6 shows the stable action of the supertip displayed in Fig. 2 after its use.

Scanning electron microscopy inspection of the deposited STM tip before and after use in the STM showed no visible changes of the supertip due to erosion or tip squeezing, which could take place during the approach or the raster motion of the tip. The tip operated successfully at both 100 and 500 mV.

If a tip 250 nm long is used employing a voltage of 1500 mV, the high resolution image of the gold crystals shows jitters at the edges of the crystals; jitters are reached first in the scan. This indicates that the emission stability changes. This effect was observed in forward and backward scanning. The jitters may also be explained by additional field emission

from sidewise located emission centers of the tip, which are activated when the tip approaches a step in the object structure.

III. CONCLUSIONS

A reproducible recipe for the fabrication of conductive supertips for STM lithography, field electron emission was obtained. Supertips built by additive lithography at high sample temperatures were characterized in a field-emission microscope and show confined emission. An ion mirror protects the cathode from ions generated at the screen and improves the stability of the emission. Deposited tips undergo a strong forming process but generate a stable field emitter with angular emission density increased to 0.2 mA/sr. Due to the use of deposited supertips on etched wires, there is no need to employ single crystalline wires as emitters in the STM. These deposited conductive tips are applicable to STM imaging. They may also work in STM lithography since they have the advantage of confining the emitted beam to one emission site delivered from a 4-nm-diam nanocrystal at the very top of the tip. Very high resolution STM lithography can be anticipated when using these tips.

ACKNOWLEDGMENTS

The authors are grateful to Dr. M. Rodewald from the Institut für Strukturforschung des Fachbereichs Materialwissenschaften der Technischen Hochschule Darmstadt for high resolution TEM images and acknowledge support by Deutsche Telekom for this work.

¹E. A. Dobisz, H. W. P. Koops, F. K. Perkins, C. R. K. Marrian, and S. L. Brandow, *J. Vac. Sci. Technol. B* **14**, 4148 (1996).

²H. W. P. Koops, R. Weiel, D. P. Kern, and T. H. Baum, *J. Vac. Sci. Technol. B* **6**, 477 (1988).

³H. W. P. Koops, J. Kretz, M. Rudolph, M. Weber, G. Dahm, and K. L. Lee, *Jpn. J. Appl. Phys.* **1** **33**, 7099 (1994).

⁴J. Kretz, Diplomarbeit, Technischen Hochschule Darmstadt, 1994; M. Rudolph, Diplomarbeit, Technischen Hochschule Darmstadt, 1994; J. Kretz, M. Rudolph, M. Weber, and H. W. P. Koops, *Microelectron. Eng.* **23**, 477 (1994).

⁵H. W. P. Koops, M. Weber, J. Urban, and C. Schössler, *Proc. SPIE* **2522**, 189 (1995).

⁶M. Takai, H. Morimoto, T. Kishimoto, S. Lipp, L. Frey, H. Ryssel, S. Yura, A. Hosono, and A. Okuda, *Technical Digest of the 9th International Conference on Vacuum Microelectronics*, 1996, p. 657.

⁷C. Schössler, A. Kaya, J. Kretz, M. Weber, and H. W. P. Koops, *Microelectron. Eng.* **30**, 471 (1996).

⁸H.-W. Fink, *J. Vac. Sci. Technol. B* **8**, 1323 (1990).

⁹H. S. Kim, M. L. Yu, U. Stauffer, L. P. Muray, D. P. Kern, and T. H. P. Chang, *J. Vac. Sci. Technol. B* **11**, 2327 (1993).

¹⁰T. H. Miller, A. Knoblauch, Ch. Wilbertz, and S. Kalbitzer, *Appl. Phys. A* **61**, 99 (1995).

¹¹A. Knoblauch, Ch. Wilbertz, Th. Miller, and S. Kalbitzer, *J. Phys. D* **29**, 470 (1996).

Simultaneous optical detection techniques, interferometry, and optical beam deflection for dynamic mode control of scanning force microscopy

M. Hoummady, E. Farnault, T. Yahiro, and H. Kawakatsu

Laboratory for Integrated Micro-Mechatronic Systems (LIMMS)/CNRS-IIS, Institute of Industrial Science,
The University of Tokyo, Roppongi 7-22-1, Minato-ku, Tokyo 106, Japan

(Received 8 September 1996; accepted 21 April 1997)

In dynamic mode control of scanning force microscopy (SFM), optical beam deflection and interferometry are the techniques most used for detection of force gradients by means of a tip and a microcantilever that usually vibrates at the first resonant mode. In order to increase the sensitivity of these kinds of microscopes, one possible means is to investigate the potential of the highest resonance modes which allow an increase in the operating frequencies. For these two detection techniques, according to the local displacement slope, care must be taken in the choice of an appropriate microcantilever point where the laser beam is focused. In this article, an original technique based on simultaneous detection, interferometry, and the beam deflection method is introduced. These techniques are able to characterize within two degrees of freedom, normal displacement and angular deflection, thus resonating the SFM cantilever. © 1997 American Vacuum Society. [S0734-211X(97)10104-4]

I. INTRODUCTION

Atomic force microscopy¹ (AFM) is a nondestructive method for imaging surface roughness of conducting and nonconducting materials with high spatial resolution. The AFM was introduced in 1986 by Binnig *et al.*, and is complementary to scanning tunneling microscopy² (STM) which is limited to conducting materials. The forces between a sample and a sharp tip are sensed using a soft cantilever spring. In dynamic mode AFM, a cantilever with the small tip at its end is driven at its resonant frequency ω_0 and scanned without contact above the surface. The resonant natural frequency is given by

$$\omega_0^2 = \frac{k_0}{m^*}, \quad (1)$$

where k_0 is the spring constant and m^* the effective mass of the cantilever. Due to tip-sample interaction at small distances (mainly van Der Waals type force in noncontact mode), the force gradients acting on the cantilever induce a shift in the resonant frequency:

$$\omega^2 = \omega_0^2 - \frac{1}{m^*} \frac{\partial F}{\partial z}. \quad (2)$$

To measure the cantilever deflection in scanning force microscopy, the optical detection techniques most used are interferometry^{3,4} and optical beam deflection⁵ (OBD). Where the first is sensitive to normal surface displacement, the second one is sensitive to cantilever angular deflection. The minimum detectable force gradient in attractive mode AFM³ varies according to $(Q\omega_0)^{-1/2}$ where Q is the quality factor. Working under a vacuum environment is one way to improve the sensitivity of the microscope thus increasing the Q factor.⁶ Otherwise, it is possible to work at high frequency driving the conventional cantilever at a higher flexural mode. We show that simultaneous interferometric and OBD measurements optimize the detection sensitivity when nodes and

antinodes appear. Moreover, simultaneous probing provides complementary information that can be useful in determining the origin of the forces acting on the cantilever.⁷

II. MECHANICAL MODEL

For our experiments we have been using a rectangular shape silicon nitride microcantilever from Olympus Optical Co., Ltd. The microcantilever has a length l of 200 μm , a width a of 20 μm , and a thickness of 8000 \AA . To calculate the resonant frequencies of flexural modes, we assume that the microcantilever is homogeneous and its cross section is uniform. The equation of motion is

$$EI \frac{\partial^4 z}{\partial x^4} + \rho A \frac{\partial^2 z}{\partial t^2} = 0, \quad (3)$$

where E , I , ρ , and A are, respectively, Young's modulus, the area moment of inertia, the density, and the cross section of the beam. A solution of the type

$$Z_n(x, t) = Z_n(x) \cos(\omega_n t + \theta) \quad (4)$$

is inserted into Eq. (3) and allows one to determine the resonant frequencies as the following:

$$\omega_n = \beta_n^2 \sqrt{\frac{EI}{\rho A}}, \quad (5)$$

where β_n is a constant depending on both the mode order n and on the length l . Taking the boundary conditions (cantilever fixed at one end and free at the other) into account, we deduce the possible values of β_n as a function of the length l :

$$\beta_1 l = 1.875, \quad \beta_2 l = 4.694, \quad \beta_3 l = 7.855,$$

$$\beta_4 l = 10.996, \dots, \quad \beta_n l = (2n-1)\pi/2.$$

Relating them to local displacement, it can be deduced from:⁸

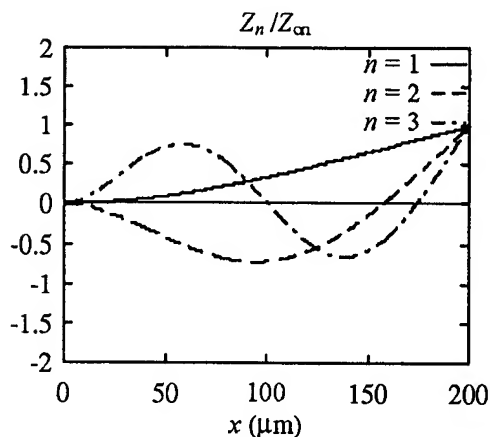


FIG. 1. Theoretical displacement for modes 1, 2, and 3. x is the position of the laser spot on the cantilever.

$$Z_n(x) = Z_{0n} [\cos(\beta_n x) - \cosh(\beta_n x) + C_n \sin(\beta_n x) - C_n \sinh(\beta_n x)], \quad (6)$$

where Z_{0n} is determined by the vibration amplitude of the mode and C_n is a factor depending on the mode order: $C_1 = -0.7341$, $C_2 = -1.0185$, $C_3 = -0.9992$, ..., $C_n = -1$.

In Fig. 1, local displacements are plotted for the three first modes. We will show in Secs. III and IV how optical methods are able to detect this kind of displacement.

III. OPTICAL BEAM DEFLECTION

Optical beam deflection was first introduced in AFM by Meyer and Amer⁵ in 1988. This technique was previously successfully implemented in photothermal spectroscopy⁹ earlier and was used widely in scanning force microscopy as a highly sensitive optical probe. A laser beam is focused onto the sample by means of a microscope objective. The reflected beam reaches a segmented photodiode that is used as a position sensitive detector (PSD). A deflection θ of the sample produces a shift in the optical power onto the PSD.¹⁰

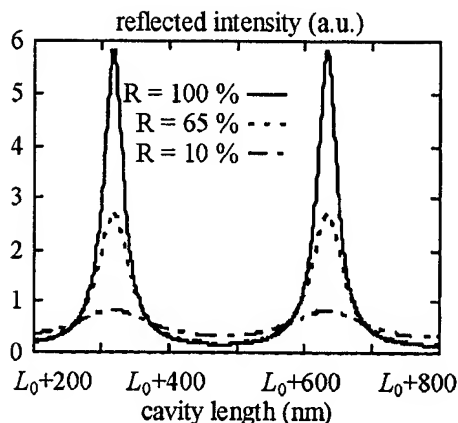


FIG. 2. Reflected intensity as a function of cavity length for the Fabry-Pérot interferometer. L_0 is the static cavity length.

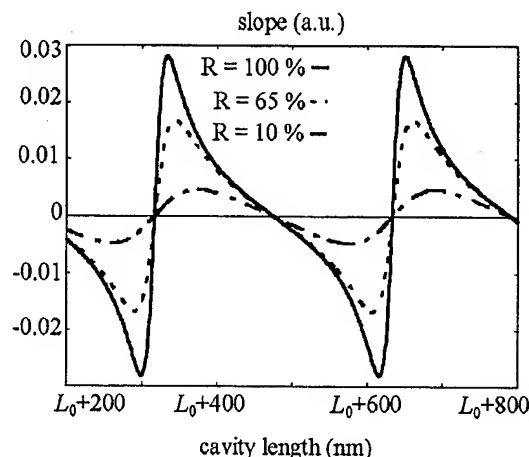


FIG. 3. The slope as a function of cavity length for the Fabry-Pérot interferometer. L_0 is the static cavity length.

$$\Delta P(x, t) = \sqrt{2\pi} \frac{P_i d_0}{\lambda} \theta(x, t), \quad (7)$$

where P_i and λ are, respectively, the incident power and the wavelength of the laser source, and d_0 is the beam diameter on the sample. Equation (7) shows that OBD is sensitive to gradient surface displacement:

$$\theta(x, t) = \frac{dZ(x, t)}{dx}. \quad (8)$$

The minimum detectable displacement for the first mode is of the order of $10^{-4} \text{ Å}/\sqrt{\text{Hz}}$. For the higher modes, the model shows that the angular deflection increases at the extremity of the cantilever. At this point, the main limitation is usually due to the driving piezoelectric transducer (PZT).

IV. INTERFEROMETRY

We have been using a Fabry-Pérot interferometer⁴ because of its high sensitivity. For this type of interferometer, the reflected intensity may be expressed as a function of the length of the cavity as follows:¹¹

$$f(x) = \frac{R(1-R_p)^2 + 2(1-R_p) \left(\sqrt{RR_p} \cos\left(\frac{4\pi x}{\lambda}\right) - RR_p \right)}{1 + RR_p - 2\sqrt{RR_p} \cos\left(\frac{4\pi x}{\lambda}\right)}, \quad (9)$$

where x is the cavity length, λ the wavelength, R_p the semi-reflecting plate coefficient of reflection, and R the sample coefficient of reflection. Figure 2 shows the reflected intensity for different values of R . The Fabry-Pérot interferometer sensitivity is optimized when the slope is at a maximum. This slope strongly depends on the sample coefficient of reflection R . In Fig. 2, $R = 100\%$ represents a perfect mirror; $R = 65\%$ is our experimental condition (a metallized cantilever with a microscope objective between the semireflecting plate and the sample). When the laser spot is focused onto a node of vibration, the relatively large deflection at this point

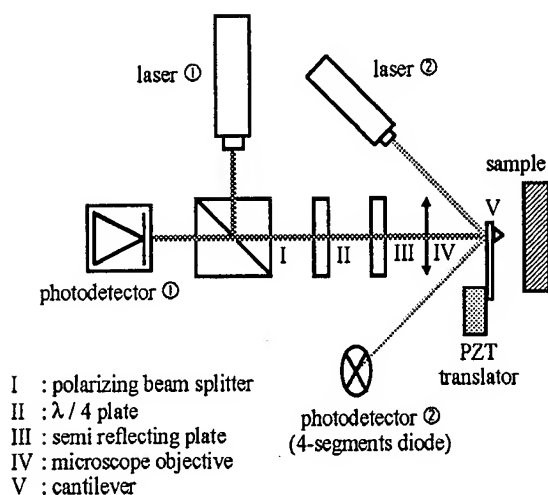


FIG. 4. Experimental setup: simultaneous interferometric and optical beam deflection detection. The Fabry-Pérot cavity is between elements III and V.

acts as a perturbation in the beam's alignment. This case corresponds to $R = 10\%$. Figure 3 shows the slope as a function of the cavity length. A low value for R corresponds to a loss of sensitivity. With ideal experimental conditions, the minimum detectable displacement using Fabry-Pérot interferometry is $10^{-5} \text{ Å}/\sqrt{\text{Hz}}$.

V. EXPERIMENTAL SETUP AND RESULTS

The experimental setup for simultaneous interferometric and deflection detection is shown in Fig. 4. The laser (1) and photodiode (1) are used for the interferometer while laser (2) and photodiode (2) are used for optical beam deflection. Figure 5 shows the experimental results obtained with the interferometric method. The laser spot is scanned along the cantilever for flexural modes 1–3 (25.6, 160.6, and 449 kHz, respectively). According to theoretical calculation, the optimized points are located on an antinode of vibration.

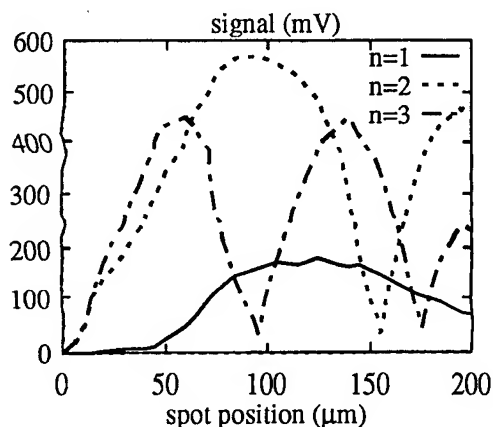


FIG. 5. Experimental displacement along the cantilever (interferometric detection) for flexural modes 1, 2, and 3.

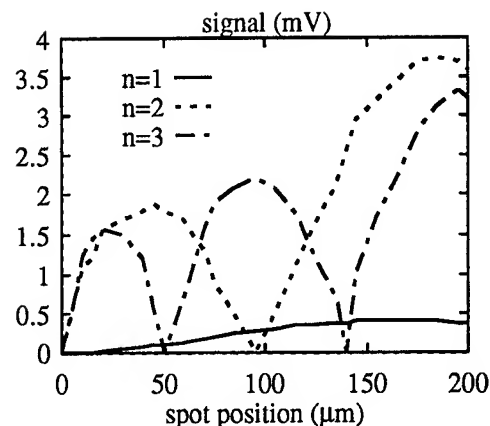


FIG. 6. Experimental angular deflection along the cantilever (OBD detection) for flexural modes 1, 2, and 3.

With the same experimental parameter, OBD measurement (Fig. 6) shows that maximum sensitivity occurs when the laser beam is focused on a node of vibration. In this case, angular deflection at the extremity of the cantilever increases with the chosen mode, improving the detection sensitivity.

In Figs. 7 and 8, the frequency is swept around the third mode (449 kHz) by means of a PZT translator, showing that the signal level is strongly altered by the position of the spot on the cantilever. The interferometric signal (Fig. 7) is obtained with the laser spot focused on the end of the cantilever and on the second antinode. In this case, the signal amplitude is divided by two. This effect is due to misalignment of incident and reflected beams. Sensitivity is then decreasing. Figure 8 shows the difference between signals measured with the OBD method for the same points. Comparing Figs. 5–8, we can deduce that the results with two optical detection are complementary. Both interferometric and OBD detection provide a method for characterizing vibrating structures and allow measurement of both displacement and angular deflection at the same time.

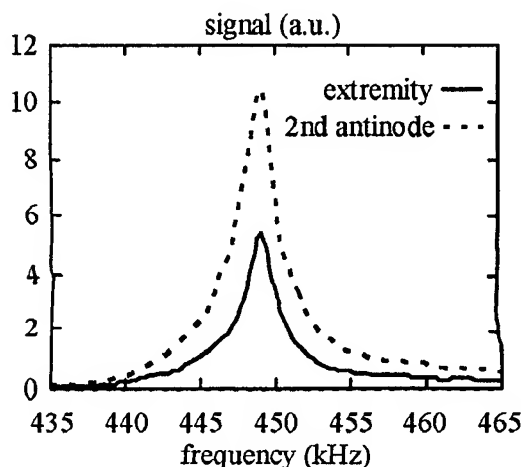


FIG. 7. Interferometric signal for mode 3 (449 kHz).

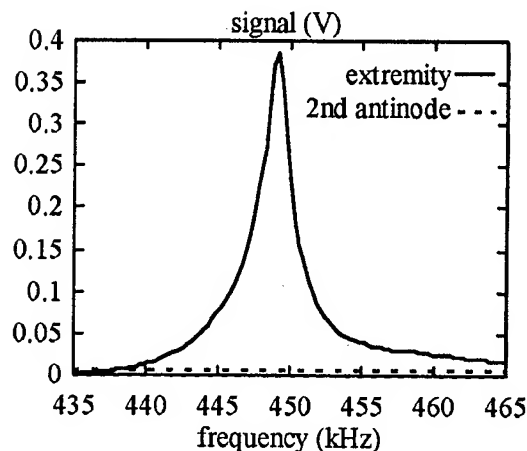


FIG. 8. OBD signal for mode 3 (449 kHz).

VI. CONCLUSION

As is well known in dynamic mode control of scanning force microscopy that the working frequency is an important factor of sensitivity for detection of gradient forces, we investigated the possibility of using a conventional cantilever (commercially available) driven at high flexural mode. An analytical mechanical model determines the position of nodes and antinodes along the cantilever. Thus experimental measurements with a combined interferometric and optical

beam deflection method are optimized according to the excited mode. With simultaneous optical methods, the displacement and displacement gradient (angular deflection) of the cantilever are characterized. This makes measurement possible within two degrees of freedom. This method provides a new tool for scanning force microscopy enhancement.

ACKNOWLEDGMENTS

The authors wish to thank Akitoshi Toda of Olympus Optical Co., Ltd. for the cantilevers that were used and Christian Pieralli and Bernard Cretin for useful discussions.

¹G. Binnig, C. F. Quate, and Ch. Gerber, *Phys. Rev. Lett.* **56**, 930 (1986).

²G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, *Phys. Rev. Lett.* **49**, 57 (1982).

³Y. Martin, C. C. Williams, and H. K. Wickramasinghe, *J. Appl. Phys.* **61**, 4723 (1987).

⁴R. Erlandsson, G. M. McClelland, C. M. Mate, and S. Chiang, *J. Vac. Sci. Technol. A* **6**, 266 (1988).

⁵G. Meyer and N. M. Amer, *Appl. Phys. Lett.* **53**, 1045 (1988).

⁶T. R. Albrecht, P. Grütter, D. Home, and D. Rugar, *J. Appl. Phys.* **69**, 668 (1991).

⁷S. Fujisawa, M. Ohta, T. Konishi, Y. Sugawara, and S. Morita, *Rev. Sci. Instrum.* **65**, 664 (1994).

⁸C. M. Harris and C. E. Crede, *Shock and Vibration Handbook* (McGraw-Hill, New York, 1976).

⁹W. B. Jackson, N. M. Amer, A. C. Boccara, and D. Fournier, *Appl. Opt.* **20**, 1333 (1981).

¹⁰J. W. Wagner, *Physical Acoustics* (Academic, New York, 1990), Vol. XIX, p. 201.

¹¹M. Françon, *Optical Interferometry* (Academic, New York, 1966), p. 108.

Development of ultrahigh vacuum-atomic force microscopy with frequency modulation detection and its application to electrostatic force measurement

Takayuki Uchihashi^{a)}

Department of Physics, Faculty of Science, Hiroshima University, Higashi-Hiroshima, Hiroshima 739, Japan

Masahiro Ohta and Yasuhiro Sugawara

Department of Electronic Engineering, Faculty of Engineering, Osaka University, Suita, Osaka 565, Japan

Yoshio Yanase

Silicon R&D Center, Sumitomo SiTix Company, Kohoku-cho, Kishima-gun, Saga 849-05, Japan

Tatsuhiko Sigematsu

Silicon Engineering Control Center, Sumitomo SiTix Company, Kohoku-cho, Kishima-gun, Saga 849-05, Japan

Mineharu Suzuki

NTT Advanced Technology Company, Atsugi, Kanagawa 243-01, Japan

Seizo Morita

Department of Electronic Engineering, Faculty of Engineering, Osaka University, Suita, Osaka 565, Japan

(Received 8 September 1996; accepted 21 April 1997)

We succeeded in high resolution force measurements by using a noncontact ultrahigh vacuum-atomic force microscope (UHV-AFM) with frequency modulation (FM) detection. We clearly observed adatoms and corner holes on the Si(111)7×7 reconstructed surface. Then we applied the noncontact UHV-AFM with FM detection to the high resolution measurement of the electrostatic force. We prevented deterioration of the spatial resolution of the topography by isolating the electrostatic interaction from van der Waals interaction. By simultaneous measurements of the topography and electrostatic force on a silicon oxide, a spatial resolution ~15 Å of the electrostatic force was achieved. © 1997 American Vacuum Society. [S0734-211X(97)10204-9]

I. INTRODUCTION

Atomic force microscopy (AFM) operating under noncontact mode is demonstrated to detect weak forces such as van der Waals,^{1,2} electrostatic,^{3,4} and magnetic forces^{5,6} on the sample surface. Noncontact AFM has an advantage in that it avoids destruction of the sample surface and the tip because of the noncontact. There are two operating modes that detect the weak force between the tip and the sample. One is the static mode which can directly detect cantilever deflection with a displacement sensor.⁷ The other is the dynamic mode using a mechanical oscillation of the cantilever. In the dynamic mode, there are two modes, one is the nonresonant mode,⁸ which measures the cantilever amplitude at lower enough frequency than the mechanical resonant frequency of the cantilever by applying force modulation between the tip and the sample. The other is the resonant mode, which measures frequency shift of the mechanical resonant frequency of the cantilever that is due to the force gradient between the tip and the sample.⁹⁻¹¹

At the dynamic-resonant mode, the sensitivity of force detection is increased dramatically because the high Q value of the cantilever reduces the thermal noise effect relative for

the cantilever oscillation.⁹⁻¹¹ On the other hand, at the static mode and the dynamic-nonresonant mode, there is no enhancement effect on the sensitivity of the force detection such as that due to the high Q value.

In the dynamic-resonant mode, a higher Q value results in higher sensitivity. In practice, the sensitivity of force detection increases in a high vacuum, because the Q value becomes higher by about several orders in high vacuum without air damping than it does in air.^{9,12,13} In addition, the spatial resolution can be improved by decreasing the separation between the tip and the sample surface.¹⁴ This small separation was realized by using a stiffer cantilever which suppresses any jumping of the tip onto the sample surface.

We were recently successful in measuring a weak van der Waals force with high sensitivity and high spatial resolution on various surfaces with ultrahigh vacuum-AFM (UHV-AFM) operating under the noncontact mode by using frequency modulation (FM) detection.^{2,15-18} Specifically, the atomic structure and atomic scale point defects were observed with true atomic resolution on a Si(111)7×7 reconstructed surface^{2,15} and on an InP(100) cleaved surface.^{16,17} It is now expected that even the electrostatic force can be measured with higher sensitivity and better resolution by using noncontact UHV-AFM with FM detection.

In this article, we first present atomic resolution measure-

^{a)}Electronic mail: uchi@ele.eng.osaka-u.ac.jp

ments on a Si(111)7×7 reconstructed surface by noncontact UHV-AFM with FM detection. Then application of the high resolution measurement of the electrostatic force on a silicon oxide is demonstrated.

II. METHOD OF SIMULTANEOUS MEASUREMENTS OF VAN DER WAALS AND ELECTROSTATIC INTERACTIONS

We assume that the conductive tip is positioned at a distance Z above an insulating film on a conductive substrate. When an ac bias voltage $V_{ac} \sin(\omega t)$ is applied between the conductive tip and the substrate, the force gradient F' acting on the tip is described by the sum of the van der Waals interaction and the electrostatic interaction. Using a parallel capacitor approximation,⁴ force gradient F' is expressed by the following:

$$\begin{aligned}
F' &= \left\{ \frac{Q_s^2}{2\pi\epsilon_0 Z^3} + \frac{1}{4} \left(\frac{\partial^2 C}{\partial Z^2} \right) V_{\text{ac}}^2 + F'_{\text{vdw}} \right\} \\
&\quad + \frac{CQ_s}{2\pi\epsilon_0 Z^2} V_{\text{ac}} \sin(\omega t) - \frac{1}{4} \left(\frac{\partial^2 C}{\partial Z^2} \right) V_{\text{ac}}^2 \cos(2\omega t) \\
&= F'_{\text{image}} + F'_{\text{cap}} + F'_{\text{vdw}} + F'_{\omega} + F'_{2\omega}. \tag{1}
\end{aligned}$$

Here, Q_s are the charges on the insulating film, C is the tip-to-sample capacitance, and ϵ_0 is the permittivity in a vacuum. The dc components in Eq. (1) are composed of the electrostatic interaction F'_{image} due to the image charge $-Q_s$ on the tip, the electrostatic interaction F'_{cap} due to capacitance (capacitive force), and the van der Waals interaction F'_{vdw} . The ω component F'_{ω} originates from the electrostatic interaction between charges Q_s on the surface and bias tip. The 2ω component $F'_{2\omega}$ originates from electrostatic interaction due to the capacitance.

In FM detection, the dc component is used to control the tip-to-sample distance. Without applying the ac bias voltage, the dc component contains F'_{vdw} and F'_{image} , whereas with applying the ac bias voltage, it is a mixture of F'_{vdw} , F'_{image} , and F'_{cap} . Atomic resolution images by noncontact UHV-AFM with FM detection utilize the F'_{vdw} because of its stronger distance dependence compared with the others. By applying the ac bias voltage, it prevents accurate control of the tip-to-sample distance Z and causes an error in the measured topography.¹⁸ F'_{image} is experimentally negligible in comparison with F'_{vdw} . So, in order to measure the topography with high resolution, subtraction of F'_{cap} in the dc component in Eq. (1) is required. We can see from Eq. (1) that F'_{cap} is equal to $F'_{2\omega}$ except for $\cos(2\omega t)$. Therefore, by subtracting the amplitude of the 2ω component from the dc component, F'_{vdw} contributes mainly to the dc component.

III. EXPERIMENT

Figure 1 shows the experimental setup for simultaneous measurements of the topography and electrostatic force by the noncontact UHV-AFM with FM detection. Force gradient F' acting on the tip was measured by the frequency shift $\Delta\nu$ of the cantilever. The cantilever was oscillated by the

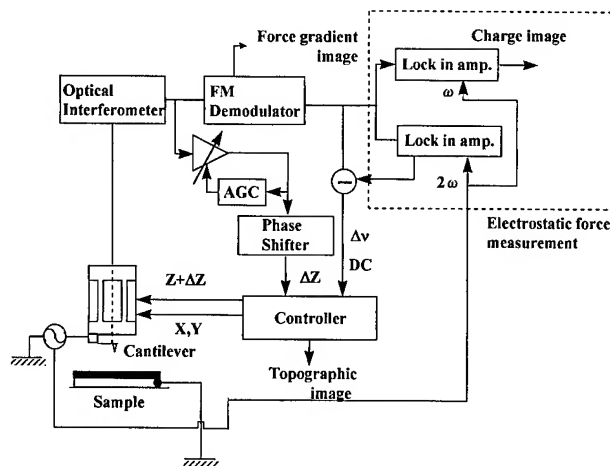


FIG. 1. Schematic diagram of the simultaneous measurement of the topography and the electrostatic force by noncontact UHV-AFM with FM detection. The part surrounded by the dotted line is used at electrostatic measurement.

piezoelectric tube scanner at the mechanical resonant frequency ν_0 under a positive feedback condition. Here, in order to avoid an unexpected fluctuation of the vibrating amplitude, a constant voltage was applied to the piezoelectric scanner using an autogain control (AGC) circuit.¹⁵ The oscillation of the cantilever was detected by fiber-optical interferometer. The frequency shifts $\Delta\nu$ of the cantilever were detected by a tunable FM demodulator.

The ac bias voltage $V_{ac} \sin(\omega t)$ was applied between the conductive cantilever and the electrically grounded substrate, and the amplitudes of the ω and 2ω components of the force gradient F' were measured by lock-in amplifiers. Here, topography was taken by using the dc component in Eq. (1), in which the amplitudes of the 2ω component F'_{cap} were subtracted as was described in Sec. II. The charge image was measured by the amplitude of the ω component F'_{ω} .

Conductive silicon cantilevers were used as force sensors. Their spring constant and mechanical resonant frequency were $k=34\text{--}36\text{ N/m}$ and $\nu_0=167\text{--}169\text{ kHz}$, respectively. The nominal radius of the tip curvature was $5\text{--}10\text{ nm}$. The Q factor of the cantilever was estimated to be about 38 000 in an UHV environment. Since the native silicon oxide layer on the Si tip surface was not removed, the surface of the tip should be covered with a nonconductive native silicon oxide layer.

Noncontact UHV-AFM measurements were performed at room temperature under a pressure lower than $2\text{--}4 \times 10^{-10}$ Torr. To measure the atomic resolution image of the Si(111) 7×7 reconstructed surface, the Si(111) sample was annealed at about 1200 °C by direct current flow for surface cleaning. On the other hand, for the electrostatic force measurement, a thin silicon oxide with an oxide thickness ~ 70 Å formed on a single crystal Si(111) wafer was used. The silicon oxide sample was placed on the sample holder in air and then it was transferred into vacuum.

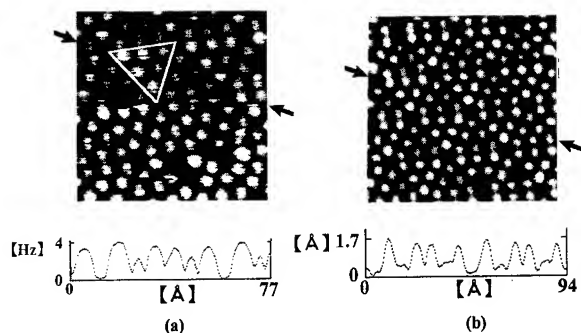


FIG. 2. Noncontact UHV-AFM images of the Si(111) 7×7 reconstructed surface and cross sectional profiles indicated by arrows. The vibration amplitude of the cantilever A_0 is 164 Å. (a) The variable force gradient image. The frequency shift $\Delta\nu$ is -17 Hz and the scan area is $72 \text{ Å} \times 72 \text{ Å}$. (b) The constant force gradient image. $\Delta\nu = -28 \text{ Hz}$ and the scan area is $89 \text{ Å} \times 89 \text{ Å}$.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Atomic resolution images of the Si(111) 7×7 reconstructed surface

The Si(111) 7×7 reconstructed surface has 19 dangling bonds in one unit cell and is very reactive.¹⁹ So far a few groups have succeeded in obtaining atomic resolution images on Si(111) 7×7 reconstructed surfaces with noncontact UHV-AFM with FM detection.^{2,15} But stable imaging has not been achieved. So we tried measurements of the Si(111) 7×7 reconstructed surface. In this measurement, it does not utilize the ac bias voltage or the part surrounded by the dotted line in Fig. 1.

Figure 2 shows atomic resolution images of the 7×7 structure of Si(111) obtained by noncontact UHV-AFM. Figure 2(a) was taken under the variable frequency shift mode. The mean frequency shift was set as $\langle \Delta\nu \rangle = -17 \text{ Hz}$ and the vibration amplitude A_0 was about 164 Å. On the other hand, Fig. 2(b) was taken under the constant frequency shift mode. The frequency shift was set as $\Delta\nu = -28 \text{ Hz}$ and the vibration amplitude A_0 was about 164 Å. In both Figs. 2(a) and 2(b), the adatoms and the corner holes according to the dimer-adatom-stacking fault (DAS) model were clearly resolved.¹⁹ The contrast in the images is similar to that of a well-known filled-state scanning tunneling microscopy (STM) image. Furthermore, atomic scale defects where adatoms were missing were observed. In Fig. 2(a), the image contrast suddenly changes halfway on the slow scans. Since the position of the adatoms is slightly shifted at that spot, this contrast change seems to be caused by the sudden change of the atom position on the tip apex.

Half of the unit cell surrounded by white triangle is brighter than the other half in Fig. 2(a). In Fig. 2(b), the same contrast can be also seen. There is a contrast difference between the halves with faults and the halves without faults. Similar results were found by Nakagiri et al.²⁰ Since we did not perform STM measurement simultaneously, we could not identify the halves we observed both with and without faults. Further, from the cross sections of Figs. 2(a) and 2(b), we can see that the height of the corner adatoms is slightly

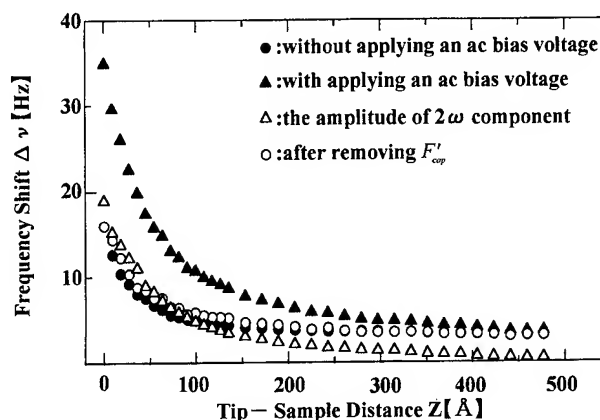


FIG. 3. The frequency shift $\Delta\nu$ as a function of the tip-to-sample distance Z on the Si(111) surface. (●), (▲) The distance dependence of the frequency shift $\Delta\nu$ with and without the ac voltage, respectively. Amplitude V_{ac} and frequency $\omega/2\pi$ of the ac bias voltage are $V_{ac} = 1 \text{ V}$ and $\omega/2\pi = 300 \text{ Hz}$, respectively. (△) The 2ω component. (○) The frequency shift $\Delta\nu$ after subtraction of the 2ω component (△) from the frequency shift (▲) at $V_{ac} = 1 \text{ V}$.

higher than that of the center adatoms. These phenomenon indicates the possibility that the images of the noncontact AFM include not only information on surface corrugation but also other local physical information, although the reason for this contrast is not yet clear.

B. Novel electrostatic force measurement with FM detection

1. Separation of the capacitive force from the distance regulation

Figure 3 shows the frequency shift $\Delta\nu$ as a function of the tip-to-sample distance Z on the Si(111) surface. Upward and downward movements on the vertical axis of the frequency shift $\Delta\nu$ correspond to an increase and decrease in the attractive force gradient. $Z=0$ was defined at the position where the vibration amplitude of the cantilever began to be unstable. The closed circles correspond to the distance dependence of the frequency shift $\Delta\nu$ at an ac bias voltage of $V_{ac} = 0 \text{ V}$. The closed triangles correspond to that at $V_{ac} = 1 \text{ V}$ and frequency $\omega/2\pi = 300 \text{ Hz}$ to the cantilever. At the same tip-to-sample distance Z , the frequency shift $\Delta\nu$ at $V_{ac} = 1 \text{ V}$ is larger than that at $V_{ac} = 0 \text{ V}$ because of the contribution from the capacitive force F'_{cap} in Eq. (1). The open triangles correspond to the 2ω component, which corresponds to $F'_{2\omega}$. The open circles correspond to the frequency shift $\Delta\nu$ after subtracting the open triangles from closed triangles at $V_{ac} = 1 \text{ V}$. The variation in the open circles is almost coincident with that of the closed circles. This result suggests that the separation F'_{cap} from the dc component in Eq. (1) is realized.

2. Simultaneous imaging of the topography and the electrostatic force distribution of SiO_2

Figures 4(a) and 4(b) show simultaneously obtained images of the topography and the ω component F'_{ω} , respec-

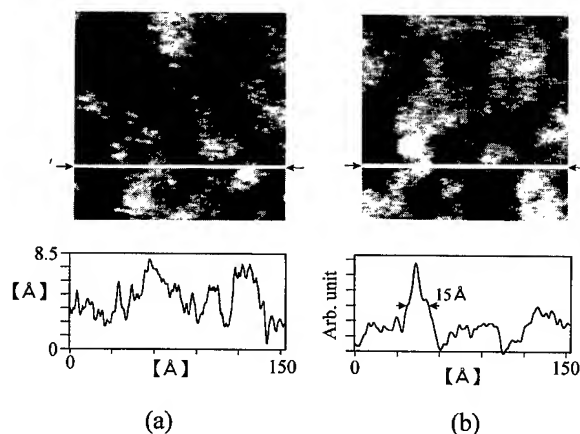


FIG. 4. (a) Topographic and (b) electrostatic force (from the ω component) images obtained simultaneously and cross sectional profiles on the position indicated by white lines. $V_{ac}=2$ V. $\omega/2\pi=1$ kHz. The scan area is $150 \text{ Å} \times 150 \text{ Å}$.

tively. A Silicon oxide with thickness $\sim 70 \text{ Å}$ was used as the sample. The scan area was $150 \text{ Å} \times 150 \text{ Å}$. From Fig. 4(a), we can see that the corrugation amplitude of the topography is less than 8.5 Å . We can see the contrast clearly on the ω component image in Fig. 4(b). We conclude that the ω component image of Fig. 4(b) is not caused by the change of the tip-to-sample distance due to the topography and that it shows electrostatic force distribution for the following reasons. First, the ω component was reproducible. Second, as seen from the cross sectional profiles, the contrast of the ω component image in Fig. 4(b) is not synchronized with that of the topography in Fig. 4(a). The lateral resolution of the electrostatic force image in Fig. 4(b) reaches to about 15 Å . It should be noted that the spatial resolution is improved one order of magnitude in comparison with the conventional electrostatic force measurement.²¹ The contrast in Fig. 4(b) seems to be caused by the electric charges trapped in silicon oxide,²² but this must be investigated in more detail in the future.

V. CONCLUSIONS

We have developed noncontact UHV-AFM with FM detection. We demonstrated atomic resolution images on a Si(111) 7×7 reconstructed surface and a high resolution electrostatic force image on the silicon oxide surface. On the Si(111) 7×7 reconstructed surface, we could clearly observe the adatoms, corner holes, and defects missing adatoms. Other interesting phenomena such as the contrast change corresponding to the difference between halves with faults and halves without faults and between corner and center ada-

toms. For the electrostatic force measurement, we proposed a novel method based on FM detection that separates the van der Waals force from the electrostatic force and prevents deterioration of the spatial resolution of the electrostatic measurement. Using this method, we obtained the topographic and the electrostatic force images simultaneously on a silicon oxide surface. From the result, the lateral resolution of the electrostatic force measurement reached about 15 Å .

ACKNOWLEDGMENTS

The authors would like to thank S. Mishima of Olympus Optical Co., Ltd. for construction of the AFM unit. A part of this work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan. The author wishes to acknowledge the Yoshida Foundation for Science and Technology.

- ¹F. Ohnensorge and G. Binnig, *Science* **260**, 1451 (1993).
- ²F. J. Giessible, *Science* **267**, 68 (1995).
- ³Y. Martin, C. C. Williams, and H. K. Wickramasinghe, *J. Appl. Phys.* **61**, 4723 (1987).
- ⁴B. D. Terris, J. E. Stern, D. Rugar, and H. J. Mammin, *Phys. Rev. Lett.* **63**, 2669 (1989).
- ⁵Y. Martin and H. K. Wickramasinghe, *Appl. Phys. Lett.* **50**, 1455 (1987).
- ⁶C. Shönberger and S. F. Alvarado, *Phys. Rev. Lett.* **65**, 3162 (1990).
- ⁷S. Morita, Y. Fukano, T. Uchihashi, T. Okusako, Y. Sugawara, Y. Yamanishi, and T. Oasa, *Jpn. J. Appl. Phys.* **32**, L1701 (1993).
- ⁸H. Yokoyama, K. Saito, and T. Inoue, *Mol. Electron. Bioelectron.* **3**, 79 (1992).
- ⁹T. R. Albrecht, P. Grütter, D. Horne, and D. Rugar, *J. Appl. Phys.* **69**, 668 (1991).
- ¹⁰R. Lüthi, E. Meyer, L. Howald, D. Anselmetti, M. Dreier, H. Rüetschi, T. Bonner, R. M. Overney, J. Frommer, and H.-J. Güntherodt, *J. Vac. Sci. Technol. B* **12**, 1673 (1994).
- ¹¹R. Lüthi, H. Haefke, K.-P. Meyer, E. Meyer, L. Howald, and H.-J. Güntherodt, *J. Appl. Phys.* **74**, 7461 (1993).
- ¹²A. Kikukawa, S. Hosaka, and R. Imura, *Appl. Phys. Lett.* **66**, 3510 (1995).
- ¹³A. Kikukawa, S. Hosaka, and R. Imura, *Rev. Sci. Instrum.* **67**, 1463 (1996).
- ¹⁴H. Rohrer, in *Scanning Tunneling Microscopy and Related Methods* (Kluwer Academic, Dordrecht, 1990), p. 1.
- ¹⁵S. Kitamura and M. Iwatsuki, *Jpn. J. Appl. Phys.* **35**, L668 (1996).
- ¹⁶Y. Sugawara, M. Ohta, H. Ueyama, and S. Morita, *Science* **270**, 1646 (1995).
- ¹⁷H. Ueyama, M. Ohta, Y. Sugawara, and S. Morita, *Jpn. J. Appl. Phys.* **34**, L1086 (1995).
- ¹⁸M. Yasutake, *Jpn. J. Appl. Phys.* **34**, 3403 (1995).
- ¹⁹K. Takayanagi, Y. Tanishiro, M. Takahashi, and S. Takahashi, *J. Vac. Sci. Technol. A* **3**, 1502 (1985).
- ²⁰N. Nakagiri, M. Suzuki, K. Okiguchi, and H. Sugimura, *Surf. Sci. Lett.* **373**, L329 (1997). 5-9-1 Tokodai, Tsukuba 300-26, Japan (private communication).
- ²¹J. Itoh, Y. Tohma, T. Inoue, H. Yokoyama, and K. Shimizu, *Jpn. J. Appl. Phys.* **33**, 7167 (1994).
- ²²Y. Fukano, T. Uchihashi, T. Okusako, K. Hontani, A. Chayahara, Y. Sugawara, Y. Yamanishi, T. Oasa, and S. Morita, *Proceedings of the International Conference on Advanced Microelectronic Device and Processing, Sendai, 1994* (Tohoku University Press, Sendai, 1994), p. 365.

Development of a metal patterned cantilever for scanning capacitance microscopy and its application to the observation of semiconductor devices

Takuma Yamamoto,^{a)} Yoshihiko Suzuki,^{b)} Masayuki Miyashita,^{c)} Hiroyuki Sugimura, and Nobuyuki Nakagiri

Tsukuba Research Laboratory, Nikon Corporation, 5-9-1 Tokodai, Tsukuba, Ibaraki 300-26, Japan

(Received 12 September 1996; accepted 31 March 1997)

The performance of scanning capacitance microscopy (SCM) strongly depends on the probe used. We have developed an original probe suitable for SCM (SCM probe), which uses a pyramid-shaped metal tip and metal lead line patterned on a silicon nitride cantilever. We installed the SCM probe on a SCM based on a commercial atomic force microscope (AFM). Differences in silicon oxide thickness and differences of dopant types and densities in a silicon substrate with a thermal oxide layer were successfully imaged by SCM simultaneously with AFM using the SCM probe. Boundaries between different dopant types and densities, which were not recognizable by AFM, were clearly observed by SCM. Signal-to-noise ratio and reproducibility were improved in the SCM images obtained with the SCM probe when compared with images obtained with a metal-coated silicon nitride cantilever. © 1997 American Vacuum Society. [S0734-211X(97)04604-0]

I. INTRODUCTION

Scanning capacitance microscopy (SCM),¹⁻³ which has been demonstrated as a tool for observing local capacitance, has the potential to be the most effective analysis technique for measuring the properties of nanometer-size semiconductor devices, such as dopant density or defects in an insulator layer, since it can nondestructively observe inner structures with high spatial resolution.

Several applications of SCM to semiconductors have been reported. Dopant profiling of cross-sectioned silicon samples was reported by Huang and Williams⁴ and by Neubauer *et al.*⁵ Their SCM data was inverted to dopant profiles and compared with profiles obtained by process simulation and secondary-ion-mass spectroscopy (SIMS). Silicon *p-n* junctions and gallium-arsenide *p-n* junctions were observed from above by Kopanski *et al.*⁶ The application of SCM to nitride-oxide-silicon (NOS) memory was demonstrated for the first time by Barrett and Quate.⁷ Further investigations of this application were reported by Dreyer and Wiesendanger.⁸ We have observed a buried structure comprised of metal lines covered with an insulator on a silicon substrate and demonstrated that SCM can detect the failure of such buried metal lines and gates in semiconductor devices.⁹ We have also demonstrated the ability of SCM to recognize differences in silicon oxide thickness.¹⁰ As indicated by these applications, SCM would be highly effective as an observation tool for semiconductor devices, the scale of which is ever shrinking.

In our previous system, the SCM signal was unstable due to damage, which occurred easily to the metal at the tip of its completely metal-coated silicon nitride probe. In order to improve the performance of SCM, a better probe satisfying

the following four requirements is demanded. First, it must have low stray capacitance with its surroundings, especially with the sample. A change in the amount of stray capacitance shifts the resonance frequency of the sensor and changes the level of the SCM signal. Any change of stray capacitance due to tip position change would appear in the SCM images. A part of the ac modulation voltage to the sample transmits to the probe due to the stray capacitance between them. This voltage then travels to the capacitance measurement sensor and offsets the SCM signal. A change in this offset signal due to generator fluctuation or for any other reason would increase the noise level. Thus, a reduction of the stray capacitance between the probe and the sample is desirable. Second, the structure of the tip end must be stable, as any damage here will degrade the final SCM image. Third, the probe must function to accurately trace the sample surface. Simultaneous atomic force microscope (AFM) imaging is the most desirable, since the force feedback of the AFM functions as a means to trace the sample surface at a constant force. Finally, the probe should be producible by a microfabrication process such as that used to make the silicon nitride cantilever widely used in AFM.

An entirely metal-coated probe does not satisfy the first and second requirements. Since both the cantilever and the probe's body are entirely coated with metal, stray capacitance is large. Furthermore, the metal at the tip end is often damaged during SCM measurement and, consequently, nonconductive silicon nitride might appear here. This is a major reason why the SCM signal is unstable. A mechanically bent, etched tungsten wire cantilever, which has sometimes been used in the past, has the advantages of small stray capacitance and tip end stability.⁴ However, it is not compatible with a commercial AFM and it is difficult to consistently fabricate many tungsten probes with the same qualities.

In order to overcome these problems, we have developed a novel metal-patterned microcantilever for SCM using a microfabrication process. We have termed it the "SCM

^{a)}Electronic mail: yamamoto@tsukubagw.nikon.co.jp

^{b)}Also at Optomechatronics R&D Dept., Nikon Co., 1-6-3 Nishi-ohi, Shinagawa, Tokyo 140, Japan.

^{c)}SPM Promotion Section, Nikon Co., 471 Nagaodai-machi, Sakae-ku, Yokohama 244, Japan.

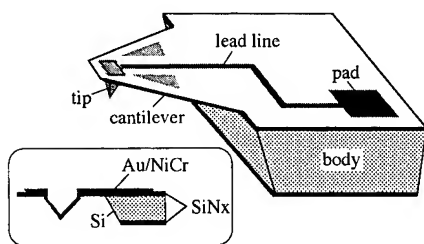


FIG. 1. Schematic diagrams of the SCM probe.

probe." The SCM probe has a patterned metal lead line and pad in order to reduce stray capacitance. Its tip is entirely made of metal to solve the problem of damage encountered with a coated-metal tip.

II. DEVELOPMENT OF THE SCM PROBE

Figure 1 shows a schematic diagram. The SCM probe consists of a silicon body, a nonconductive thin film which forms a triangular cantilever, a metal lead line and pad, and a metal tip. Since the tip is entirely made of metal, the tip end can maintain conductivity in spite of a certain amount of damage. The pad on the body, which is electrically connected to the tip through the lead line patterned on the cantilever, is used for the electrical connection to the capacitance sensor circuit. The cantilever is $680\text{ }\mu\text{m}$ long and its base is $700\text{ }\mu\text{m}$ wide. Its force constant is estimated at 0.0014 N/m . Its measured resonant frequency is 1.6 kHz . The lead line is $6\text{ }\mu\text{m}$ wide at the cantilever and $40\text{ }\mu\text{m}$ at the body. The pad is $300\text{ }\mu\text{m} \times 400\text{ }\mu\text{m}$. We also coated the regions on both sides of the lead line on the cantilever with metal in order to reflect the AFM laser beam. These regions are electrically independent from the metal lead line at the center.

Figure 2 shows the fabrication process for the SCM probe. A 3-in. (100) oriented *n*-type silicon wafer of $250\text{ }\mu\text{m}$ thickness was used. First, silicon nitride (SiN_x) film was de-

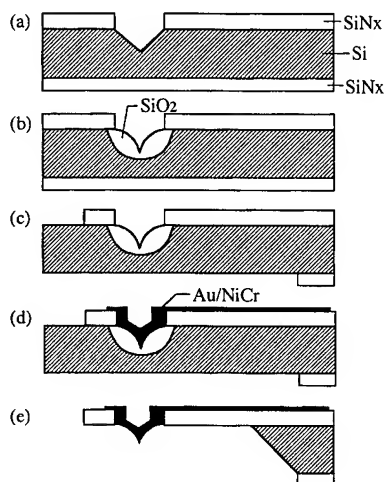


FIG. 2. Fabrication of the SCM probe.

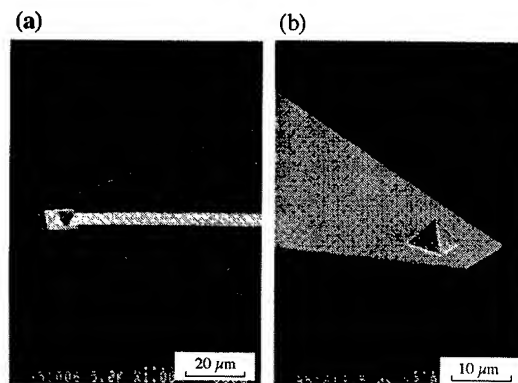


FIG. 3. Scanning electron micrographs of the SCM probe from the top (a) and from the bottom, i.e., tip side (b).

posited to a thickness of $0.7\text{ }\mu\text{m}$ by low-pressure chemical vapor deposition (LPCVD). Next, the film was patterned photolithographically to form $6\text{ }\mu\text{m}$ square openings for microtip patterns. The exposed silicon was etched with an anisotropic etchant, tetramethyl ammonium hydroxide [TMAH; $(\text{CH}_3)_4\text{NOH}$], to form pyramidal micropits [Fig. 2(a)]. After these micropits were formed, the exposed silicon was oxidized by thermal oxidation [Fig. 2(b)]. The wafer was then patterned and aligned so that the end of each cantilever was positioned around a single micropit. The SiN_x film on the bottom of the wafer was patterned to define the shape of the body [Fig. 2(c)]. Next, spin-coated photoresist film was patterned photolithographically to form the pattern for the metal tip, lead line, and pad. First, nickel chromium (NiCr) 2 nm thick, and then gold (Au) 200 nm thick were deposited by the electron beam evaporation method. Unnecessary resist film having a metal layer was then removed in acetone [Fig. 2(d)]. Finally, in order to form the free-standing cantilever, the wafer was dipped in potassium hydroxide [KOH] aqueous solution to etch away the exposed silicon [Fig. 2(e)].

Figure 3 shows scanning electron micrographs of the SCM probe. The radius of the pyramidal tip was estimated to be less than 50 nm . Figure 3(a) was taken from the top side of the cantilever and Fig. 3(b) was taken from the bottom, that is, tip side. The SCM probe shown here does not have the metal-coated regions designed to reflect the AFM laser beam, as described above. We fabricated versions of the SCM probe both with the beam reflection regions and without, in order to judge the effect of the stray capacitance between the lead line and the reflection regions. However, no such effect was observed and we therefore took simultaneous SCM and AFM images using the SCM probe with the beam reflection regions.

III. SCM OBSERVATION WITH THE SCM PROBE

Figure 4 shows a schematic diagram of our SCM combined with an AFM. We constructed our SCM by adding a capacitance sensor and other electronic circuits to a commercial AFM (CP, Park Scientific Instruments). An ac modulation voltage and a variable dc bias voltage were applied to

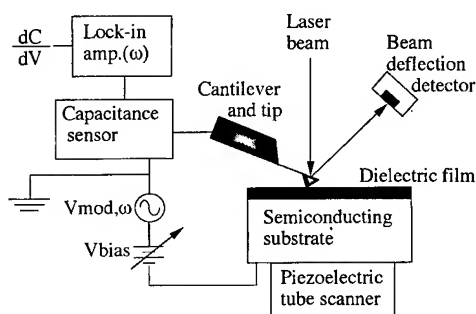


FIG. 4. Schematic diagram of the scanning capacitance microscope.

the sample which was placed upon a piezotube scanner. The ac component of the modulation frequency in the output signal (dC/dV) was detected using a lock-in amplifier and used as the SCM data.

In SCM, the capacitance between the conductive tip and the sample is detected by a capacitance sensor which was originally developed by RCA as a detective circuit for their Video Disk.¹¹ The tip is electrically connected to the LC resonator in the sensor. Tip-sample capacitance change shifts the resonant frequency of the resonator. As a result, the capacitance or its change is detected by the sensor. We used this same type of sensor, in our case, one which had been manufactured for JVC's Video Disk system. The capacitance of a metal-oxide-silicon (MOS) structure constructed by the metal tip of the SCM and a semiconductor sample with an insulator layer is much smaller than that of an ordinary MOS structure with a metal gate. Therefore, SCM measures not the capacitance itself but derivatives of it, mostly to eliminate the effect of stray capacitance.

We imaged two types of samples by SCM. Figure 5(a) is a cross section of the first sample, which consisted of silicon oxide locations of different thickness on silicon. A (100)-

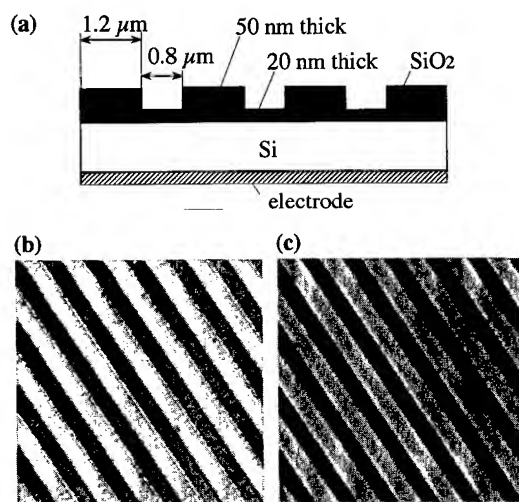


FIG. 5. (a) Cross section of a sample having an oxide layer of different thicknesses. (b) AFM image taken with the SCM probe. (c) Simultaneously acquired SCM image. In the SCM image, the data was adjusted so that the maximum and minimum values were normalized to 255 and 0, respectively.

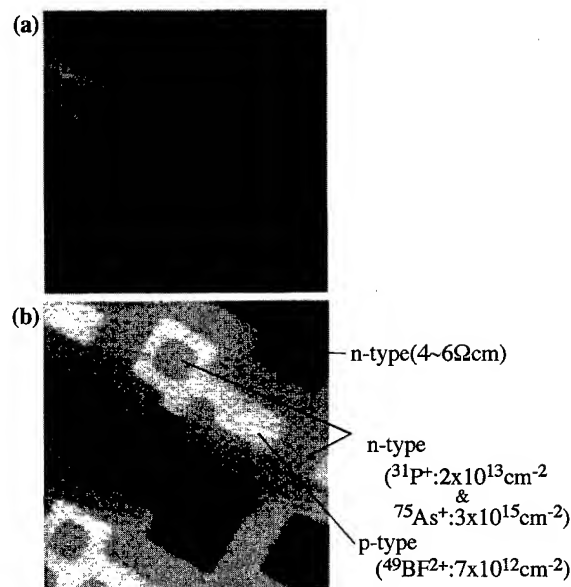


FIG. 6. Simultaneously obtained AFM image (a) and SCM image (b) of a sample with different dopant types and densities. In the SCM image, the data was adjusted so that the maximum and minimum values were normalized to 255 and 0, respectively.

oriented 4–6 Ω cm *n*-type Si wafer was cleaned by the conventional RCA method and thermally oxidized at 900 °C (H_2/O_2 : 19 min 20 s; O_2 : 1 min; N_2 : 2 min) to grow a 50-nm-thick SiO_2 layer. A 1- μ m-thick aluminum layer was deposited by vacuum evaporation on the bottom side of the Si wafer to act as an electrode. A resist pattern of lines and spaces was made on the oxide layer, after which the oxide was partly etched by 1% HF for 6 min. The sample was then washed in acetone to remove the resist. Consequently, a line-shaped structure was fabricated on the oxide. Figure 5(b) is an AFM image of the sample. The oxide thickness difference of 30 nm was clearly imaged by the SCM probe. Figure 5(c) is a SCM image of the same sample taken simultaneously with the AFM image. A modulation voltage (80 kHz, 4 Vp p) was applied to the sample. The dc bias was 0 V in this case. The scan area was 12 μ m \times 12 μ m. The AFM image and the SCM image in Fig. 5 are consistent. The low (i.e., thin) oxide regions are dark in the SCM image due to the large negative dC/dV value of the *n*-type silicon substrate.

The second sample was an *n*-type silicon wafer of 4–6 Ω cm which had regions of different doping. Simultaneously taken SCM and AFM images of the sample are shown in Fig. 6. A donor, P^+ was partly implanted at a dose of 2×10^{13} atoms/cm² at 100 keV. Silicon nitride was used as a mask in this process. As^+ was then partly implanted at a dose of 3×10^{15} atoms/cm² at 120 keV and annealed (1000 °C, 30 min) to form high density *n*-type regions. *p*-type regions were made by partly implanting BF_2^+ at a dose of 7×10^{12} atoms/cm² at 100 keV and annealed at 950 °C for 30 min. No additional dopant was implanted at the darkest areas in Fig. 6(b). The surface of the sample was entirely covered with a thermal oxide layer. As shown in Fig. 6(b), differences of dopant types and densities of silicon were success-

fully imaged by the SCM. For example, the contrast between the *n*- and *p*-type regions, which cannot be seen in the AFM image [Fig. 6(a)], is clearly recognizable in the SCM image. The SCM image in Fig. 6(b) was taken with an 80 kHz, 4 Vp-p modulation voltage and 0 V dc bias. The scan area was $20\text{ }\mu\text{m}\times 20\text{ }\mu\text{m}$. In this image, the BF^{2+} implanted *p*-type regions were bright due to their positive dC/dV values and the no-dose *n*-type regions were dark due to their negative dC/dV values. The dC/dV signal of the high density *n*-type regions was not detectable.

It took about 5 min to take the simultaneous SCM and AFM images shown in Figs. 5 and 6. The time constant of the lock-in amplifier was 3 ms in these cases. The longer the time constant is, the better the signal-to-noise (S/N) ratio becomes. However, the amount of time necessary to acquire the SCM image increases proportionately.

IV. DISCUSSION

In our previous experiments employing an entirely metal-coated silicon nitride probe,^{9,10} the capacitance between the tip and the sample was in the order of 10^{-18} F. The capacitance between the lever and the sample was about 10^{-17} F and that between the body and the sample was more than 2×10^{-14} F. Since the capacitance is proportional to the area of the metal coating, stray capacitance can be effectively reduced by not coating the entire lever and body with metal, and instead employing a patterned metal lead line and pad. The total capacitance between the new SCM probe and a sample is roughly estimated to be 2×10^{-15} F. This value is about one-fiftieth of the estimated stray capacitance between an entirely metal-coated probe and sample. However, in our experiment, the amplitude of the undesired signal detected by the capacitance sensor due to the transmission of ac modulation voltage from the sample to the probe through the capacitance between them did not reduce. This unwanted signal had the same frequency as, but a different phase from, the dC/dV signal. The fact that the stray capacitance did not reduce was probably due to the conductivity of the silicon body. The silicon body formed capacitances of more than 10^{-13} F both with the pad and with the sample.

Although the stray capacitance did not reduce as expected, the reproducibility and the S/N ratio of the SCM were unmistakably improved by the SCM probe most likely due to the stability of its entirely metal tip. Previously, when using an entirely metal-coated cantilever for SCM/AFM, the coated metal often peeled off at the tip end. This was probably due to the weak adhesion of the coated metal to the end of the sharp tip. Residual stress in the coated metal at the tip might cause such peeling off when the tip contacted and scanned a sample surface. The SCM probe does not suffer from such a problem, since its tip is entirely made of metal.

Several changes would improve the performance of the SCM probe. First, the body should be made of an insulator in

order to achieve a reduction of stray capacitance with the sample. A glass body as used in a conventional silicon nitride cantilever would be one candidate. Second, the dimensions of the cantilever should be optimized. The SCM probe's cantilever is very long compared with commercially available cantilevers and is not ideal for AFM due to its low resonant frequency. A shorter, thicker cantilever would provide a higher resonant frequency. Third, other materials should be considered for the patterned metal. It's likely that there is a more suitable pattern material for the SCM probe than gold. Harder metals would more strongly resist deformation at the tip end. Finally, a metal tip covered with an insulator layer would be necessary to measure a sample without an insulator layer. The oxide layer of a metal, like TiO or Al_2O_3 , might be suitable as a stable insulator layer for such applications.

V. SUMMARY

We have developed the SCM probe which uses a pyramid-shaped metal tip and metal lead line patterned on a silicon nitride cantilever and body. We installed it on our SCM and conducted simultaneous SCM/AFM imaging of two different samples. Differences in silicon oxide layer thickness and in the dopant types and densities of silicon were successfully recognized. Stray capacitance between the SCM probe and the sample did not reduce as expected, probably due to the conductivity of the silicon body. However, contrast and repeatability improved probably due to the SCM probe's entirely metal tip. Several improvements would increase both the capabilities and applications of the SCM probe.

ACKNOWLEDGMENTS

The authors would like to thank A. Kamashita for preparing the second sample, M. Ishikawa for fabrication of the mechanical holder, and R. Motoori for his help with the electronics.

¹J. R. Matey and J. Blanc, *J. Appl. Phys.* **47**, 1437 (1985).

²C. D. Bugg and P. J. King, *J. Phys. E* **21**, 147 (1988).

³Y. Martin, D. W. Abraham, and H. K. Wickramasinghe, *Appl. Phys. Lett.* **52**, 1103 (1988).

⁴Y. Huang, C. C. Williams, and J. Slinkman, *Appl. Phys. Lett.* **66**, 344 (1995).

⁵G. Neubauer, A. Erickson, C. C. Williams, J. J. Kopanski, M. Rodgers, and D. Adderton, *J. Vac. Sci. Technol. B* **14**, 426 (1996).

⁶J. J. Kopanski, J. F. Marchiando, and J. R. Lowney, *J. Vac. Sci. Technol. B* **14**, 242 (1996).

⁷R. C. Barrett and C. F. Quate, *J. Appl. Phys.* **70**, 2725 (1991).

⁸M. Dreyer and R. Wiesendanger, *Appl. Phys. A* **61**, 357 (1995).

⁹N. Nakagiri, T. Yamamoto, H. Sugimura, and Y. Suzuki, *J. Vac. Sci. Technol. B* **14**, 887 (1996).

¹⁰T. Yamamoto, Y. Suzuki, H. Sugimura, and N. Nakagiri, *Jpn. J. Appl. Phys.* **35**, 3793 (1996).

¹¹R. C. Palmer, E. J. Denlinger, and H. Kawamoto, *RCA Rev.* **43**, 194 (1982).

Novel high vacuum scanning force microscope using a piezoelectric cantilever and the phase detection method

Jiaru Chu,^{a)} Toshihiro Itoh, Chengkuo Lee,^{b)} and Tadatomo Suga

Research Center for Advanced Science and Technology, The University of Tokyo, Komaba 4-6-1, Meguro-ku, Tokyo 153, Japan

Kazutoshi Watanabe

Seiko Instruments Incorporated, Oyama-cho, Sunto-gun, Shizuoka 410, Japan

(Received 12 September 1996; accepted 31 March 1997)

A novel high vacuum scanning force microscope (HV-SFM) using a piezoelectric cantilever and the phase detection method was developed. A self-excited force sensing PZT microcantilever with dimensions of $125 \times 50 \times 3.58 \mu\text{m}$ is vibrated in vacuum at its resonance frequency by applying an ac voltage to its PZT layer. The piezoelectric microcantilever possesses a spring constant of 16.7 N/m and a vibrational quality factor of 300 in air and 807 in vacuum. As we know, force gradients acting on the vibrating cantilever cause changes in vibration amplitude and phase. The change of the phase of the output piezoelectric current, which corresponds to the phase variation of cantilever vibration, is measured as the controlling signal in this method. The piezoelectric HV-SFM that was constructed has been proved to be stable and easy to handle in vacuum. Its vertical resolution is higher than 1.1 Å. High resolution images of an evaporated Au film can be obtained in vacuum by this new dynamic HV-SFM. Grains of about 5 nm can be seen clearly. The theoretical sensitivity of phase detection with piezoelectric cantilevers is also analyzed and compared with that of amplitude slope detection in this article. © 1997 American Vacuum Society. [S0734-211X(97)03604-4]

I. INTRODUCTION

The scanning force microscope (SFM) has been widely used for observing a variety of samples, including insulators which cannot be imaged in a scanning tunneling microscope (STM). Most of the SFMs currently in use are only capable of observing samples in air. Because measurement is performed in ambient conditions, the sample surface is covered with a thin layer of water. Due to the adhesive force of the water layer, it is rather difficult to control the loading force on the sample surface. In order to reduce the loading force, the dynamic mode of operation, which includes the cyclic-contact mode and the noncontact mode, is used for organic and biological samples to prevent surface damage.

The main reason that a vacuum SFM has not become a standard technique is that the structures of the force sensing units in traditional SFMs are too complicated. Generally, a SFM measures forces by determining the deflection of a cantilever beam. The cantilever deflection is sensed by electron tunneling, interferometry, or optical beam bouncing. In all these deflection sensing methods, the deflection detection elements must be precisely aligned to the cantilevers. These elements are too complex and cumbersome to be operated in a vacuum chamber. Another problem for the most commonly used optical beam bouncing method is that in vacuum the energy of the laser beam propagating on the end of the can-

tilever will heat the cantilever and cause higher thermal noise.

So force sensing microcantilevers, such as the piezoresistive type,¹⁻³ the piezoelectric type,⁴⁻⁶ and the capacitive type,⁷ have been proposed to construct scanning force microscopes without an external deflection sensor. Among those sensing schemes, the piezoelectric scheme is the most promising one for constructing a compact and easily operating dynamic HV-SFM because the piezoelectric cantilever can be excited directly by an imposed ac voltage without any disturbance of the sensing action or extra noise from the external oscillator. The capacitive cantilever can be excited as well; however, the resonance frequency of the cantilever is affected by the imposed ac voltage. The first type, the piezoresistive cantilever, only possesses a sensing function. Furthermore, the piezoelectric microcantilever has even been proved to be able to work as a Z scanner.⁸

The noncontact mode of the dynamic SFM is based on the measurement of the cantilever resonance shift due to the force gradient acting on the tip. In the methods used to detect the force gradient in the noncontact SFM, slope detection in ambient is the most commonly used one due to its simple control logic.^{9,10} Frequency modulation detection is also attractive because it can work well in ultrahigh vacuum (UHV) with a cantilever of high mechanical quality (typically a factor higher than 10^4 for a commercial Si_3N_4 cantilever in UHV) and can keep away from low frequency disturbance.^{11,12} This article will show that the phase detection method is a good choice when the cantilever has a quality factor in the range that gives acceptable responding speed (for example, piezoelectric cantilevers that have typical quality factors no higher than 3000 in the high vacuum condi-

^{a)}Permanent address: Department of Precision Engineering, University of Science and Technology of China, Hefei 230026, China.

^{b)}Also at: Mechanical Engineering Laboratory, AIST, Ministry of International Trade and Industry, Namiki 1-2, Tsukuba, Ibaraki 305, Japan; Present address: Microsystems Laboratory, ITRI, Bldg. 52, 195 Sec. 4, Chung Hsing Rd., Chutung, Hsinchu 310, Taiwan.

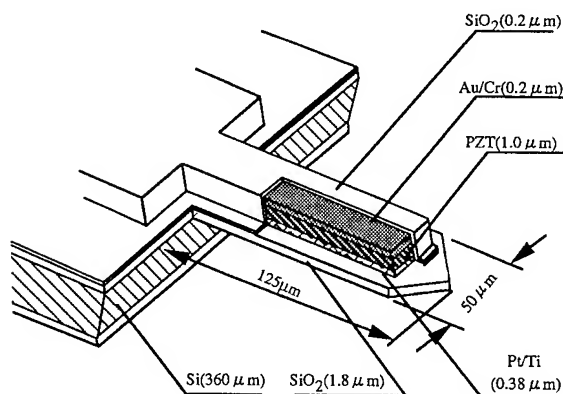


FIG. 1. Schematic drawing of a piezoelectric cantilever.

tion), because phase detection can give higher sensitivity than amplitude slope detection with a simple control logic.

In our previous study, we studied the SFM using a piezoelectric cantilever for use in the ambient condition. In this article, we report the development of a noncontact SFM using a self-excited PZT microcantilever for use in high vacuum, in which a newly developed phase detection method for cantilever vibration sensing and feedback control is used.

II. HIGH-VACUUM SFM USING A PIEZOELECTRIC CANTILEVER AND THE PHASE DETECTION METHOD

A. Structure and microfabrication of the piezoelectric cantilever

The piezoelectric cantilever developed consists of a 1.8- μm -thick thermal SiO_2 layer, a 1.0- μm -thick piezoelectric layer sandwiched between 0.2- μm -thick Au/Cr and 0.38- μm -thick Pt/Ti electrodes, and a 0.2- μm -thick passivation layer of SiO_2 . The schematic drawing of the cantilever is shown in Fig. 1. The piezoelectric layer is a PZT [$\text{Pb}(\text{Ti}_{0.52}\text{Zr}_{0.48})\text{O}_3$] thin film that was fabricated by sol-gel method. Details of the PZT thin film preparation are described elsewhere.⁶ In the micromachining processes, the metal and PZT layers were patterned by an Ar ion beam. The reactive ion beam of C_3F_8 and O_2 is used for etching the SiO_2 layer. Finally, freestanding cantilevers are formed by removing the underneath bulk silicon by immersion in KOH solution. The dimensions of the levers are $125 \times 50 \times 3.58 \mu\text{m}$. In order to offset capacitance current and resistance current from the cantilever current, a reference pattern with the same area as the cantilever is prepared on the die of the chip. A scanning electron microscopy (SEM) micrograph of 125- μm -long piezoelectric cantilevers is shown in Fig. 2. Anyone of the cantilevers can be used for force microscopy. Table I shows the characteristics of the cantilevers fabricated. The spring constant is the value calculated from the cantilever dimensions and layer materials with the formula described by Itoh and Suga.^{13,14}



FIG. 2. SEM photograph of the piezoelectric cantilevers.

B. Instrumentation of the novel HV-SFM system

When we excite the cantilever near its resonance frequency, the phase of the cantilever vibration can be expressed as a function of the ratio of the driving frequency and the resonance frequency of the cantilever. For a large space between tip and sample, where the effect of the sample on the tip can be neglected, the phase of the tip vibration Φ can be written as¹¹

$$\Phi = \tan^{-1} \left(\frac{r}{Q(1-r^2)} \right), \quad (1)$$

where $r = \omega_d / \omega_0$, ω_0 is the resonance frequency of the cantilever during free vibration, and ω_d is the driving frequency. Clearly, in order to get the biggest change in phase of lever vibration for a given change in resonance frequency $\Delta\omega_0$, one would work on the steepest portion of the Φ vs r curve [Eq. (1)]. With the limit of large Q , the maximum slope of the Φ vs r curve is found to be at $r = 1 - 1/(8Q^2)$, as opposed to $r = 1 \pm 1/(8Q)$ in amplitude slope detection, and at this point on the curve

$$\frac{\partial \Phi}{\partial r} \approx 2Q \quad (2)$$

when a small force gradient F' ($F' \ll k$) acts on the free end of the lever, the resonance frequency will be shifted by $d\omega_0 = \omega_0 F' / (2k)$, and thus cause a phase change $\Delta\Phi$ given by

$$\Delta\Phi = - \frac{QF'}{k}; \quad (3)$$

TABLE I. Characteristics of the piezoelectric microcantilever for the HV-SFM.

Lever length	Resonance frequency	Spring constant ^a	Quality factor (in vacuum)
125 μm	163.6 kHz	16.7 N/m	807
Sensitivity 0.76 nA/nm	Capacitance 56 pF	Resistance of the PZT layer 46 M Ω	

^aCalculated value.

Eq. (3) means that the force gradient acting on the oscillating cantilever can be sensed by detecting the phase of cantilever vibration.

In order to excite the cantilever and detect its displacement with the piezoelectric cantilever, ac voltage is applied directly to the piezoelectric layer and the displacement is detected by measuring the current produced. The piezoelectric charge current reflects the mechanical resonance property near the lever resonance frequency.⁵ When the ac voltage $V = V_0 \cos(\omega t)$ (ω is the driving frequency; t is the time) is applied to the piezoelectric layer of the uniform unimorph cantilever, the output charge induced by the piezoelectric effect is given by⁵

$$q_c = \frac{d_{31} Y_p (y_p + h_p/2) w}{K_f} \int_0^1 M(x) dx, \quad (4)$$

where w is the width of the lever; d_{31} is the piezoelectric strain constant of the length expansion; Y_p and h_p are Young's modulus and the thickness of the piezoelectric layer, respectively; y_p is the distance between the plane of the zero strain and the bottom of the piezoelectric layer; K_f is the flexural rigidity of the lever; and $M(x)$ is the bending moment which is determined by the vibration mode. The current from the piezoelectric layer, which is dependent on the piezoelectric charge output, the capacitance C , and the resistance R of the layer, is expressed by the following:

$$Y = \frac{dq_c}{dt} \cos(\omega t + \Phi_c) + C\omega \cos(\omega t + 90^\circ) + \frac{V}{R} \cos(\omega t), \quad (5)$$

then the phase of the current from the cantilever is given by

$$\Phi = \tan^{-1} \frac{(dq_c/dt) \sin \Phi_c + C\omega V}{(dq_c/dt) \cos \Phi_c + (V/R)}, \quad (6a)$$

where Φ_c is the phase of the piezoelectric charge current, which is the reflection of the phase of the cantilever vibration. The changing range of Φ_c is 180° . But because of the existence of C and R , the range of the detected phase Φ is reduced. In the case of a $1.0\text{-}\mu\text{m}$ -thick PZT layer the resistance R and the capacitance C were measured as $46\text{ M}\Omega$ and 56 pF , respectively. In this case, the range of phase Φ detected is only about half that of Φ_c and the sensitivity will be greatly affected. In order to solve this problem, we designed a PZT film (reference pattern) with the same area as the cantilever on the die to subtract the C and R current from the cantilever current. After that, Eq. (6a) becomes

$$\Phi = \Phi_c. \quad (6b)$$

The influence of capacitance and resistance of the cantilever on the detected phase signal can thus be eliminated and the detected phase reflects the phase of the cantilever vibration.

Figure 3 is a schematic diagram of the dynamic HV-SFM using a PZT microcantilever and the phase detection method. The excitation ac voltage signal is applied to the cantilever and to the reference pattern. The piezoelectric charge current, which was obtained by subtracting the current of the reference pattern from the current of the cantilever, is used as the

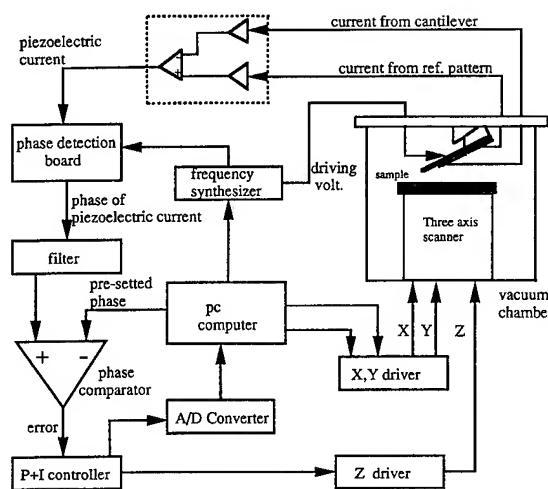


FIG. 3. Schematic diagram of the HV-SFM using the self-excited PZT cantilever and the phase detection method. The excitation voltage is applied to the PZT layer and to the reference pattern on the die. The phase of the vibration of the cantilever is used as the feedback signal to keep the distance between the cantilever and the sample surface constant during X and Y scanning.

input signal of the phase detection circuits. The phase detection circuit was built with a phase-locked loop IC CD54HCT7046A (Harris Semiconductor). That is an exclusive-OR network. Its characteristic is $V_{\Delta\Phi} = (V_{cc}/\pi) \times (\Phi_{\text{sign}} - \Phi_{\text{ref}})$, where $V_{\Delta\Phi}$ is the output voltage, V_{cc} is the supply voltage, Φ_{sign} and Φ_{ref} are the phases of input signal and reference signal, respectively. The converting ratio of the phase detection circuit is 55.5 mV/deg . The phase measured from the above phase detection circuit is compared with the phase of the set point set by the computer through a digital/analog (D/A) converter. The error signal is amplified and sent to the scanner for actuating in the Z axis to keep the phase of the cantilever vibration at the set point during X - Y scanning. The xyz scanner is a tube-type piezo scanner (scanning area xy : $20\text{ }\mu\text{m}$; z : $2\text{ }\mu\text{m}$), which is placed in the vacuum chamber.

C. Microscopy with the HV-SFM constructed

Figure 4 shows the amplitude and the phase spectra of the piezoelectric current taken in vacuum of $6.5 \times 10^{-7}\text{ Torr}$ with $125\text{-}\mu\text{m}$ -long cantilever. The applied ac voltage was kept at 13 mV during frequency scanning. Comparing the phase curve and the amplitude curve in Fig. 4, we can find that the noise obtained from the phase signal is much smaller than that from the amplitude signal when the piezoelectric charge current is as small as several nA. The signal to noise ratio of the former is estimated as two times of the latter. This may be attributed to the fact that low frequency disturbance, such as temperature change, environmental vibration, and power supply fluctuation, will superimpose noises to the output current signal, whereas the measured phase signal of the output current will be free of those influences.

Figure 5 shows a SFM image obtained by the HV-SFM using a piezoelectric cantilever with the phase detection method. The sample is an evaporated Au film that was de-

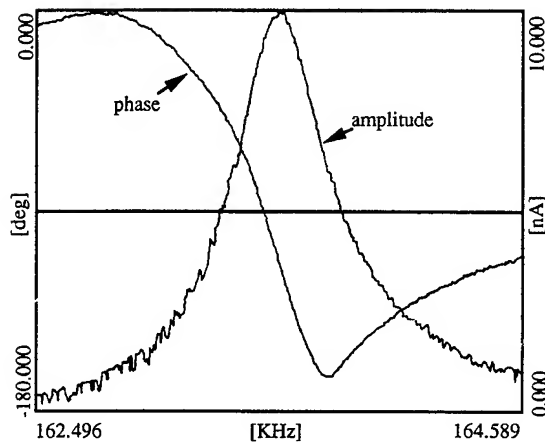


FIG. 4. Phase and amplitude spectra of the piezoelectric current for the 125- μm -long PZT microcantilever. The ac voltage applied to the PZT thin film layer was kept constant during frequency scanning.

posited on the Si substrate at 100 °C. During scanning of this image, the cantilever is driven at its natural frequency. The free vibrational amplitude was estimated as 22 nm. The morphology of undulatory agglomerated grains as small as 5 nm can be seen clearly. The vertical resolution of the instrument can be derived from the slope of the phase versus tip-sample separation curve (force curve) and the noise of the phase signal; they were measured as 3.5 deg/1 nm and 20 mV, respectively. The vertical resolution is calculated as about 1.1 Å.

III. MINIMUM DETECTABLE FORCE GRADIENT WITH THE PHASE DETECTION METHOD

The vertical resolution of the noncontact SFM can be rephrased by the minimum detectable force gradient. In the conventional noncontact SFM system with an external displacement sensor, thermal vibration of the cantilever limits the minimum detectable frequency shift under most conditions. The thermal noise¹² excites the lever with a vibration noise $N = \sqrt{4K_B T Q B / k \omega_d}$, at a frequency ω_d and in bandwidth B . The phase measurement uncertainty due to thermal

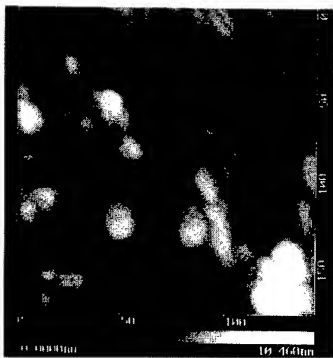


FIG. 5. An image of the evaporated Au film on the silicon substrate taken with the HV-SFM using a piezoelectric cantilever and the phase feedback method. Small grains of 5–10 nm can be seen clearly.

noise is $\Delta\Phi_n = N/A_0$, where A_0 is the amplitude of the cantilever vibration. The smallest detectable force gradient F' is found by setting $\Delta\Phi = \Delta\Phi_n$, where $\Delta\Phi$ is expressed in Eq. (3) as

$$F'_{\text{min-th}} = \sqrt{\frac{4K_B T B k}{Q \omega_0 A_0^2}}. \quad (7)$$

In the case where the SFM is constructed using the piezoelectric cantilever, $\Delta\Phi_n$ may be dependent on the electrical noise of the piezoelectric layer. The electric charge noise can be determined by the resistive Johnson noise $\langle q_{jn}^2 \rangle$, which is the mean square amplitude of the generated charge caused by Johnson noise (jn). When the minimum and maximum frequency of the bandwidth B are $\omega_1 (= \omega_d - \pi B)$ and $\omega_2 (= \omega_d + \pi B)$, respectively, $\langle q_{jn}^2 \rangle$ can be written as¹³

$$\langle q_{jn}^2 \rangle = \frac{2}{\pi} K_B T C^2 \int_{\omega_1}^{\omega_2} \frac{R}{1 + (\omega C R)^2} d\omega, \quad (8)$$

where R and C are the total resistance and the capacitance of the system, respectively, R is calculated from the resistance of the piezoelectric layer and the input impedance of the preamplifier, and C is calculated from the permittivity of the piezoelectric layer and additional capacitance due to the input cable and preamplifier. ΔZ due to $\langle q_{jn}^2 \rangle$ is given by $\Delta Z_{jn} = \langle q_{jn}^2 \rangle^{1/2} / S_0$, where S_0 is output charge amplitude per unit vibration amplitude of the lever end. We can express S_0 at the frequency ω_d by

$$S_0 = \frac{\omega_d \tau}{\sqrt{1 + \omega_d^2 \tau^2}} S_i, \quad (9)$$

where τ is a time constant of the system and S_i is the induced piezoelectric charge per unit vibration at the lever end. Then, the measurement uncertainty of the phase due to Johnson noise can be given from $\Delta\Phi_n = \Delta Z_{jn} / A_0$, Eq. (8) and Eq. (9), as

$$\Delta\Phi_{jn} = \sqrt{\frac{4K_B T B}{R S_i^2 \omega_0^2 A_0^2}}, \quad (10)$$

set to $\Delta\Phi = \Delta\Phi_{jn}$, we can get the minimum detectable force gradient due to Johnson noise from Eq. (3) and Eq. (10) as

$$F'_{\text{min-jn}} = \sqrt{\frac{4K_B T B k^2}{R Q^2 S_i^2 \omega_0^2 A_0^2}}. \quad (11)$$

Comparing Eq. (7) and Eq. (11) with the corresponding equations of amplitude slope detection method described in Ref. 15, we find that they have identical forms except for the coefficients. The coefficient in Eqs. (7) and (11) are smaller than those in the corresponding equations in the amplitude slope detection. With all parameters same, the F'_{min} in the phase detection method is about 39% of that in the amplitude slope detection method. This is attributable to the difference between the slope of the Φ vs F' curve and the slope of the A vs F' curve. Therefore, theoretically, the sensitivity in phase detection is about 2.56 times that in amplitude slope detection.

$F'_{\min-jn}$ is dependent on system resistance R . Since the impedance of the preamplifier can be large enough, the system resistance can be approximately equal to the resistance of the piezoelectric film. Therefore, the resistance of the piezoelectric film determines which factor, thermal noise or Johnson noise, governs the F'_{\min} . From Table I, the resistivity of the PZT layer on the cantilever is calculated as $28.7 \times 10^4 \Omega m$; according to Fig. 3 in Ref. 15, F'_{\min} in this instrument is governed by thermal noise. This means that this HV-SFM has a theoretical sensitivity equal to the one in the conventional system which consists of the cantilever and an external displacement sensor.

IV. CONCLUSIONS

A novel noncontact HV-SFM using a piezoelectric microcantilever and the phase detection method was successfully developed. Because no external deflection sensing unit and no external oscillating element are needed, the structure of the HV-SFM constructed is as simple as that of a regular HV-STM. The phase detection technique can offer higher vertical resolution than amplitude slope detection because the slope of the phase versus F' curve is steeper than that of the amplitude versus F' curve and, furthermore, the phase signal is free from low frequency environmental disturbances that oppose the amplitude signal. The HV-SFM constructed has been proved to be stable and easy to be handle. Very clear images of Au film can be obtained with this instrument. The morphology of undulatory agglomerated grains as small as 5 nm can be seen clearly. The vertical resolution of this instrument is estimated as 1.1 Å.

ACKNOWLEDGMENT

The authors would like to thank Dr. R. Sawada of NTT Opto-electronics Laboratories for his assistance in the cantilever fabrication.

- ¹M. Tortonese, R. C. Barret, and C. F. Quate, *Appl. Phys. Lett.* **62**, 834 (1993).
- ²M. Tortonese, H. Yamada, R. C. Barrett, and C. F. Quate, *Proceedings of Transducers '91*, IEEE Publication No. 91 CH2817-5 (1991), p. 448.
- ³S. C. Minne, S. R. Manalis, A. Atalar, and C. F. Quate, *J. Vac. Sci. Technol. B* **14**, 2456 (1996).
- ⁴T. Itoh and T. Suga, *Nanotechnology* **4**, 218 (1993).
- ⁵T. Itoh, T. Ohashi, and T. Suga, *IEICE Tran. Electron.* **E78-C**, 146 (1995).
- ⁶C. Lee, T. Itoh, and T. Suga, *Proceedings of the IUMRS-ICEM '94 Symposium, Hsinchu, Taiwan* (Materials Research Society-Taiwan, Taiwan, 1994), p. 21.
- ⁷N. Blanc, J. Brugger, N. F. de Rooji, and U. Durig, *J. Vac. Sci. Technol. B* **14**, 901 (1996).
- ⁸T. Itoh, C. Lee, and T. Suga, *Appl. Phys. Lett.* **69**, 2036 (1996).
- ⁹Y. Martin, C. C. Williams, and H. K. Wickramasinghe, *J. Appl. Phys.* **61**, 4723 (1987).
- ¹⁰G. Meyer and N. M. Amer, *Appl. Phys. Lett.* **53**, 2400 (1988).
- ¹¹T. R. Albrecht, P. Grutter, D. Horne, and D. Rugar, *J. Appl. Phys.* **69**, 668 (1991).
- ¹²A. Kikukuwa, A. Hosaka, Y. Honda, and S. Tanaka, *Appl. Phys. Lett.* **61**, 2607 (1992).
- ¹³T. Itoh and T. Suga, *J. Micromech. Microeng.* **5**, 231 (1995).
- ¹⁴T. Itoh and T. Suga, *Sens. Actuators A* **43**, 305 (1994).
- ¹⁵T. Itoh, T. Ohashi, and T. Suga, *J. Vac. Sci. Technol. B* **14**, 1577 (1996).

New technique for nanocantilever fabrication based on local electrochemical etching: Applications to scanning force microscopy

M. Hoummady, E. Farnault, H. Fujita, H. Kawakatsu, and T. Masuzawa

Laboratory for Integrated Micro-Mechatronic Systems (LIMMS)/CNRS-IIS, Institute of Industrial Science, University of Tokyo, Minato-ku, Tokyo 106, Japan

(Received 8 September 1996; accepted 28 April 1997)

We propose a new method to fabricate a nanocantilever with a mechanical oscillating ball on its end. The method is based on local electrochemical etching in a very thin electrolyte liquid layer. It allows fabrication of a very sharp tip as well as a ball tip that can be used in dynamic mode control of atomic force microscopy. The oscillating part measures 50 nm in diameter; the diameter of the ball tip is 1.5 μm . Because of the small mass of the oscillating ball and the small size of the threadlike neck, the estimated resonance frequency is around 300 MHz. This technique will help to increase the force gradient sensitivity in dynamic mode control of atomic force microscopy. © 1997 American Vacuum Society. [S0734-211X(97)14404-3]

I. INTRODUCTION

During the last decade, the scanning tunneling microscopy (STM),¹ invented in 1982, has contributed to the development of several other scanning probe techniques, such as atomic force microscopy (AFM) in 1986,² magnetic force microscopy (MFM) in 1987,³ as well as other types of microscopy.⁴ These kinds of microscopes are currently being employed and are now considered excellent tools for mapping and profiling surfaces with atomic resolution. Moreover, they opened the door to new applications such as the manipulation of nanostructures, data storage, nanolithography, and more.

All these techniques depend greatly on the quality of the probes. In STM, a sharp metallic tip that is positioned at the tunneling distance from sample allows measurement of the tip/sample electronic interaction with high lateral and vertical resolution. The sharpness of the tip is one factor that affects the resolution. In AFM contact mode, specifications such as sharpness of the tip and spring constant of the cantilever supporting the tip are also important. At atomically close separation, interaction between the tip and sample is dominated by short range interatomic forces. In noncontact mode, it is possible to measure long range interactions, as, e.g., van der Waals forces, magnetostatic, and Coulomb forces. As reported by Hartmann,⁵ Garcia and Binh,⁶ and Binh *et al.*,⁷ a spherical tip supported by a threadlike neck is well suited for this kind of measurement.

In this article, we propose a new method to fabricate very short cantilever with spherical tips for dynamic (oscillating cantilever) noncontact AFM. According to the small size of the spherical tip and supporting threadlike neck of a few tens of nanometers, the resonant frequencies expected are in the range of 100 MHz–1 GHz. As the sensitivity of dynamic-mode control increases with higher resonant frequency,⁸ it will improve force gradient sensitivity in dynamic-mode control of the AFM.

The fabrication method is based on local electrochemical etching/polishing in a very thin electrolyte layer around 1 μm thick. This technique is suitable for fabrication of both

sharp tips as well as three-dimensional structures with convolution symmetry such as the case of cylinders, spheres, cones, etc. The fabrication method is presented in Sec. II and in Sec. III we will give a theoretical estimate of the resonant frequency of the nanocantilever with the ball tip at its end.

II. NANOCANTILEVER FABRICATION TECHNIQUE

For the preparation of the spherical tip, Garcia and Binh and Binh *et al.*^{6,7} proposed a method based on surface self-diffusion. It consists of heating a STM tip in a vacuum environment to produce mobility of the atoms towards the apex of the tip and then to form a ball at the end of the tip. A criticism of this technique is that the mobility of the atoms occurs at various regions of the tip, making fabrication very difficult to control. On the other hand, electrochemical etching techniques are well controlled and are currently used to fabricate sharp tips for STM and field ion microscopy (FIM).^{9,10} We propose extending this technique to achieve local etching/polishing in a very thin electrolyte layer around 1 μm thick.

The experimental setup is presented in Fig. 1. It consists of a thin mesh containing an electrolyte liquid that is maintained by surface tension. The wire to be etched is inserted into the mesh by means of an X-Y-Z nanotranslator stage. A microscope with a magnification factor of 5000 and a charge coupled device (CCD) camera are used to control the position of the wire in the electrolyte as well as to visually control the size of the wire during the etching procedure. For our experiments we used tungsten wire, KOH as the etching solution, and a platinum mesh as the counterelectrode. A scanning electron microscope (SEM) image of the mesh used is shown in Fig. 2. The size of each hole is $150 \times 150 \mu\text{m}^2$ in area and 4 μm thick. The size of the mesh is very important as it determines the length of the threadlike neck that supports the ball tip.

The W wire, which typically has an initial diameter of 100 μm is moved, by means of the nanotranslator stage, through one of the mesh holes containing the KOH solution. Either ac and dc voltage can be used. Generally, we start by

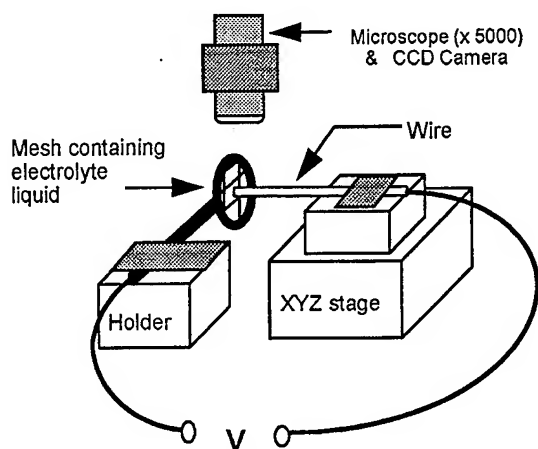


FIG. 1. Experimental setup employed for fabrication of the ball tip.

reducing the diameter of the W wire using rapid ac etching followed by a dc electropolishing to attain the shape desired. Figure 3 shows a simple procedure by which to make a ball tip. After reducing the size of the wire, the ball tip is obtained by a combination etching of the selected region and scanning the wire through the electrolyte thin film in order to obtain a smooth surface. By controlling the etching time and the region to be etched, it is possible to make any structure with cylindrical symmetry. In Fig. 4 a perfect ball tip fabricated by this technique is shown. The diameter of the ball is $1.5\ \mu\text{m}$ and the diameter of the threadlike neck is around $50\ \text{nm}$.

By periodically moving the W tip through the electrolyte film, it is also possible to make an array of structures with cylindrical symmetry such as that shown in Fig. 5. It is mainly an array of constrictions of etched threadlike necks spaced from one another by a $10\ \mu\text{m}$ length.

Due to the small size of the nanocantilever (neck) and the small mass of the oscillating ball, the natural frequency typically lies in the range of $100\ \text{MHz}$ – $1\ \text{GHz}$; it will be estimated more precisely in Sec. III.

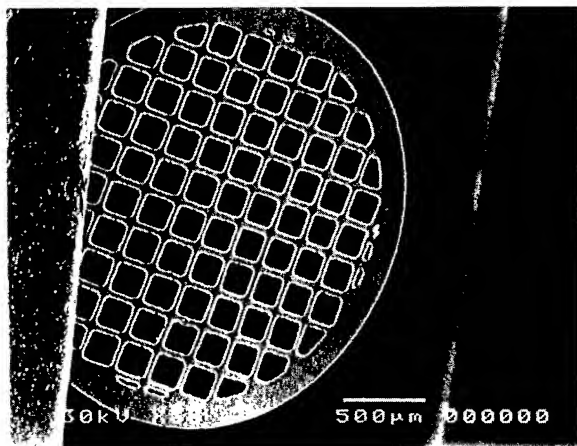


FIG. 2. SEM image showing the mesh used for local electrochemical etching.

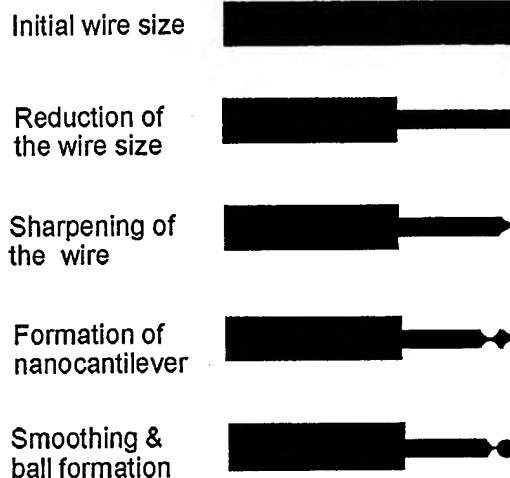


FIG. 3. Procedure by which to make a ball tip supported by a nanocantilever.

III. RESONANT FREQUENCY

In dynamic-mode AFM, the cantilever plays a key role in the measurement of long range forces. The forces acting on the resonating cantilever induce a shift in the resonant frequency. The increase of the operating resonant frequency allows the force gradient sensitivity to be improved as shown by the following equation:⁸

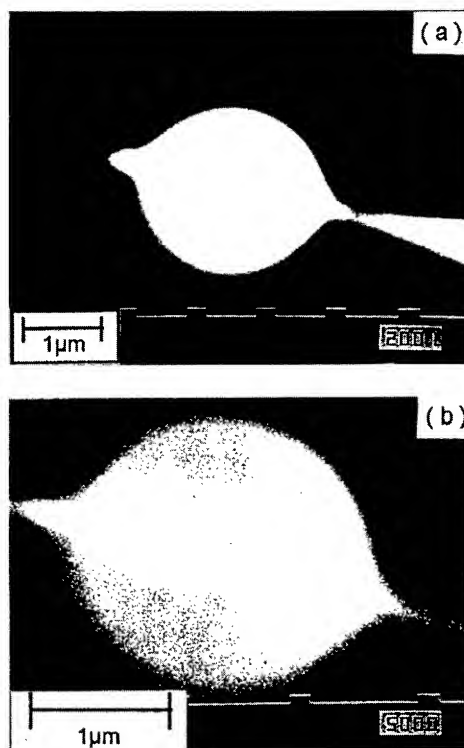


FIG. 4. SEM image of the ball tip and the nanocantilever. (a) Low magnification inspection. The diameter of the ball is $1.5\ \mu\text{m}$. (b) High magnification of the nanocantilever. The diameter is around $50\ \text{nm}$.

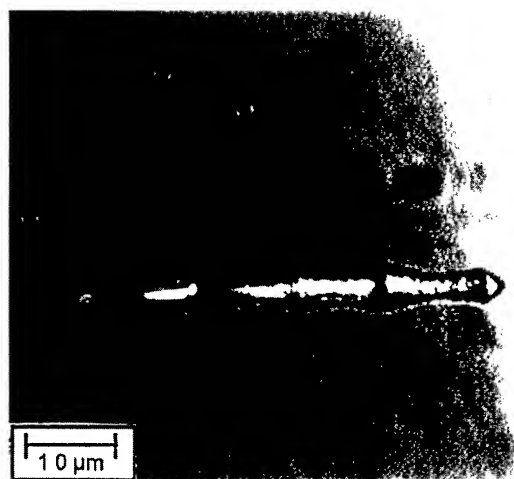


FIG. 5. Photograph of a microarray of constrictions (threadlike necks) spaced from one another by a 10 μm length.

$$F'_{\min} = \frac{1}{a_0} \sqrt{\frac{4k_0 K_B B T}{\omega Q}}, \quad (1)$$

where a_0 is the amplitude of the cantilever vibration, k_0 is the cantilever spring constant, K_B is the Boltzmann constant, T is the temperature, and B is the measurement bandwidth; ω is the cantilever frequency and Q is the resonance quality factor. Note that an AFM cantilever with small mass allows a reduction of the value of $(k_0/\omega)^{1/2}$, which is the relevant figure of merit that is to be minimized.

To calculate the resonant frequency of a nanocantilever with a mass at its end, theoretical modeling of lumped systems can be used. For either a cylindrical or a rectangular cantilever, the natural frequency is given by¹¹

$$\omega_0^2 = \frac{k_0}{m_{\text{eff}}}, \quad (2)$$

where k_0 is the spring constant and m_{eff} is the effective mass. For the first resonant flexural mode, we can define m_{eff} by

$$m_{\text{eff}} = m_C + 0.24m_d, \quad (3)$$

where m_C is concentrated mass and m_d is the distributed mass of the cantilever. To increase the resonant frequency both masses should be very small. For a cylindrical cantilever, resonant frequency is given by¹²

$$\omega = \sqrt{\frac{3\pi E r^4}{4l^3(m_C + 0.24\pi r^2 l \rho)}}, \quad (4)$$

where E , ρ , r , and l are, respectively, the Young's modulus, the density, the radius of the constriction part (threadlike neck), and its length.

To estimate the resonant frequency of ball tip supported by the nanocantilever using the following typical values of $m_C = 10^{-14}$ g, $r = 25$ nm, and $l = 200$ nm, we found that the natural resonant frequency is up to 300 MHz. In comparison to conventional resonant frequencies of an actual AFM, our probe should significantly improve the sensitivity of the AFM.

IV. CONCLUSION

A new technique for fabrication of a nanocantilever with a ball tip was reported. Based on local electrochemical etching, this technique allows realization of the both the STM sharp tip as well as the AFM ball tip that can be used to detect long range forces such as van der Waals force, etc. Moreover, due to the high resonant frequency, this kind of probe has the potential to significantly improve the sensitivity of the AFM.

ACKNOWLEDGMENTS

The authors wish to thank Professor Okano and Professor Watanabe from IIS-University of Tokyo for their fruitful discussions. This research was partially supported by research funding by the Institute of Industrial Science (University of Tokyo).

¹G. Binnig, H. Rohrer, Ch. Gerber, and E. Wiebel, *Appl. Phys. Lett.* **40**, 178 (1982).

²G. Binnig, C. F. Quate, and Ch. Gerber, *Phys. Rev. Lett.* **56**, 930 (1986).

³Y. Martin and H. K. Wickramasinghe, *Appl. Phys. Lett.* **50**, 1455 (1987).

⁴G. M. McClelland, R. Erlandson, and S. Chiang, *Review of Progress in Quantitative Nondestructive Evaluation*, edited by D. O. Thompson and D. E. Cimonti (Plenum, New York, 1987), Vol. 6B, p. 1307.

⁵U. Hartmann, *J. Vac. Sci. Technol. B* **9**, 465 (1991).

⁶N. Garcia and V. T. Binh, *Phys. Rev. B* **46**, 7946 (1992).

⁷V. T. Binh, N. Garcia, and A. L. Levanuyk, *Surf. Sci. Lett.* **301**, L224 (1994).

⁸T. R. Albrecht, P. Grütter, D. Horne, and D. Rugar, *J. Appl. Phys.* **69**, 669 (1991).

⁹J. P. Ibe et al., *J. Vac. Sci. Technol. A* **8**, 3570 (1991).

¹⁰A. J. Melmed, *J. Vac. Sci. Technol. B* **9**, 601 (1991).

¹¹*Shock and Vibration Handbook*, edited by C. M. Harris and C. E. Crede (McGraw-Hill, New York, 1976), pp. 7-13, 715.

¹²D. Sarid, *Scanning Force Microscopy* (Oxford University Press, New York, 1991), pp. 11-17.

Development of a piezoelectric self-excitation and self-detection mechanism in PZT microcantilevers for dynamic scanning force microscopy in liquid

Chengkuo Lee^{a)}

Nanometer-scale Manufacturing Science Laboratory, RCAST, The University of Tokyo, Meguro-Ku, Tokyo 153, Japan, and Department of Manufacturing Systems, Mechanical Engineering Laboratory, AIST, MITI, Tsukuba, Ibaraki 305, Japan

Toshihiro Itoh and Takahiro Ohashi^{b)}

Nanometer-scale Manufacturing Science Laboratory, RCAST, The University of Tokyo, Meguro-Ku, Tokyo 153, Japan

Ryutaro Maeda

Department of Manufacturing Systems, Mechanical Engineering Laboratory, AIST, MITI, Tsukuba, Ibaraki 305, Japan

Tadatomo Suga

Nanometer-scale Manufacturing Science Laboratory, RCAST, The University of Tokyo, Meguro-Ku, Tokyo 153, Japan

(Received 12 September 1996; accepted 4 March 1997)

We report on the first successful operation of a cyclic contact scanning force microscope (SFM) using a self-excited cantilever in liquid. Based on a new structure including a PZT reference pattern and a free-standing PZT cantilever, the piezoelectric self-excitation and self-detection mechanism for a PZT microcantilever was developed successfully. The topography is recorded by measuring the piezoelectric current variation, which corresponds to the vibration. Since the acoustic excitation from the external oscillator no longer exists, a clear single resonance peak can be obtained by using this self-excited force sensing PZT cantilever. Clear grating imaging in liquid was demonstrated, and it is compatible with the image taken in air. The future potential of applying this piezoelectric SFM to the characterization of biological samples in liquid is very promising. © 1997 American Vacuum Society. [S0734-211X(97)05904-0]

I. INTRODUCTION

Since scanning force microscopy (SFM) was introduced by Binnig *et al.* in 1986,¹ its applications in biological science have progressed rapidly.² Among those applications, imaging of DNA and living cells in solution has become one of the most importance. Many researchers have employed contact mode SFM in such applications, but the cantilever tip acting on the surface could be a major cause of possible deformation and the slight modification of soft biological samples. The use of the cyclic contact mode SFM, i.e., the so-called tapping mode SFM, for imaging soft biological samples can reduce the influence of lateral forces and incur less deformation on soft biological samples.^{3,4} In the cyclic contact mode, the cantilever is oscillated near its resonance frequency and only periodically touches the sample surface at the bottom of each swing; then the variation of vibrational amplitude is recorded as the image signals for sample topography. A microcantilever, an external piezoelectric oscillator, and an optical displacement sensing component are necessary constituents for a cyclic contact mode SFM system.

Three force sensing schemes, i.e., piezoresistive, piezoelectric, and capacitive, are applied to activate the force sensing cantilevers.⁵⁻⁷ Use of these force sensing cantilevers can eliminate the optical displacement sensor that occupies most of the space in a SFM and that also imposes operational difficulty for the SFM due to unavoidable alignment problems. An external piezoelectric oscillator will induce a severe damping influence on the cantilever when it operates in a liquid environment because of its relatively large volume. Only the self-excited force sensing microcantilever based on the piezoelectric scheme can overcome the drawbacks caused by the optical sensor and external oscillator without loss of performance and resolution in the SFM.⁸⁻¹¹ For a SFM equipped with a self-excited force sensing piezoelectric microcantilever, the essential component for cyclic contact mode SFM becomes simply the piezoelectric cantilever itself.

In this study a piezoelectric PZT microcantilever for the cyclic contact SFM in liquid is proposed. Based on the successfully developed piezoelectric self-excitation and self-detection mechanism for this PZT microcantilever, the operation of such cantilevers in liquid is characterized. Clear images of a grating sample are obtained by a SFM using this self-excited force sensing PZT microcantilever operated in a liquid environment. The images are compatible with the images taken in an ambient air environment.

^{a)}Present address: Microsystems Lab., ITRI, Bldg. 52, 195 Sec. 4, Chung Hsing Rd., Chutung, Hsinchu, Taiwan 310, R.O.C.; Electronic mail: 860004@hq.itri.org.tw

^{b)}Present address: Toto Ltd., 2-8-1, Honson, Chigasaki-city, Kanagawa Prefecture 253, Japan.

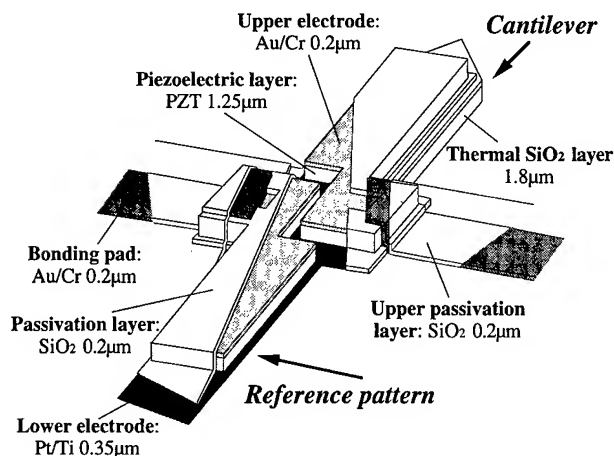


FIG. 1. Schematic drawing of the self-excited force sensing PZT microcantilever for SFM used in liquid.

II. PIEZOELECTRIC EXCITATION AND DETECTION MECHANISM

The design and structure of the self-excited force sensing PZT microcantilever are schematically illustrated in Fig. 1. The 1.25- μm -thick PZT film is deposited by a sol-gel process.¹² Ar ion beam etching is used to pattern the electrodes and the PZT layer, while reactive ion beam etching of $\text{C}_3\text{F}_8/\text{O}_2$ is applied to pattern the SiO_2 layer. Details of the microfabrication process of the self-excited force sensing PZT microcantilever are reported elsewhere.¹⁰ Characteristics of the microfabricated PZT cantilever are outlined in Table I. The piezoelectric PZT cantilever is able to be excited by an applied ac voltage via the inverse piezoelectric effect, while the force sensing is executed by recording the piezoelectric current change due to the fact that the PZT layer may give a sensitive field response to weak stress through the direct piezoelectric effect. In order to measure the piezoelectric current caused by the cantilever vibration precisely, the capacitance current output from the piezoelectric cantilever must be negated because the piezoelectric layer of the cantilever is a capacitor as well.⁸ A piezoelectric excitation and detection method for piezoelectric ZnO microcantilevers was proposed by Itoh and Suga.⁸ A reference CR circuit was used to compensate for the capacitance current from the ZnO cantilevers. For these the capacitor and resistor, which have the same values as the capacitance and resistance of the ZnO cantilevers, were chosen as the reference circuit.

TABLE I. Characteristics of the self-excited force sensing PZT microcantilever.

Lever length (160 μm)	Lever width (50 μm)	PZT layer length (135 μm)	Calculated spring constant (10.5 N/m)
Resonance frequency (in air) 112.8 kHz	Resonance frequency (in 2-propanol) 60 kHz	Quality factor (in air) 200	Quality factor (in 2-propanol) 6

The capacitance of the ZnO layer is about 1 pF, whereas the capacitance of the PZT layer is about 50 pF. The capacitance current from the PZT cantilever is higher than the capacitance current from the ZnO cantilever. In addition, the capacitance value of the PZT layer is found to be influenced by the value and frequency of applied ac voltages in the high frequency range.¹³ The capacitance value of the PZT layer also varies slightly for each different PZT cantilever. These factors make the offset canceling operation via outside electronics difficult and clear piezoelectric current signals for the PZT cantilevers become impossible to be read. As a result, in the present study the new structure includes a free-standing cantilever, and a reference pattern is developed for making the offset cancellation. Since this PZT reference pattern is the same size as the PZT cantilever, it is expected to possess the same capacitance and resistance as the cantilever. A differential current amplifier is used to record the current output coming from the reference pattern and the microcantilever when they both receive the same ac voltage. The piezoelectric current signals can then be recorded by deducting the capacitance current, which is the output from the reference pattern, from the output signals of cantilever. The variation of the vibrational amplitude can be measured by recording the piezoelectric current signals from the differential current amplifier. Figure 2(a) shows the current output versus the driving frequency from the cantilever and reference pattern when the current output from both is recorded separately. The piezoelectric current versus the driving frequency recorded from the output of a differential current amplifier is shown in Fig. 2(b). Successful offset cancellation was done easily using this cantilever-reference pattern structure. Compared to use of outside electronics for cancellation of the different piezoelectric cantilevers, the outside electronics themselves have to be modified depending on the capacitance and resistance of the different cantilevers. The differential current amplifier used here is sufficient for different cantilevers.

III. DYNAMIC PIEZOELECTRIC SFM IN LIQUID

A. Vibration versus driving frequency characteristics

For the cyclic contact mode operation in liquid, the cantilever has to be excited at its resonance. Hansma *et al.* used the piezoelectric scanner to excite the cantilever by acoustic waves through the surrounding liquid,³ while Putman *et al.* used a cantilever holder with an external piezoelectric oscillator to oscillate the cantilever.⁴ Due to the strong viscous damping effect caused by the external oscillator, the cantilever is acoustically excited rather than mechanically excited by the oscillator.⁴

In present study, in order to compare the two methods to each other, the PZT cantilever is either excited by an external oscillator or self-excited. The cyclic contact mode SFM using the self-excited force sensing PZT cantilever in liquid is schematically depicted in Fig. 3. The liquid used is 2-propanol, $\text{CH}_3\text{CH}(\text{OH})\text{CH}_3$, which has a boiling point of 82.3 $^\circ\text{C}$ and a density of 0.78 g/cm^3 (from Wako Chemicals, Co., Ltd., Tokyo). In Fig. 3 the PZT cantilever is self-excited

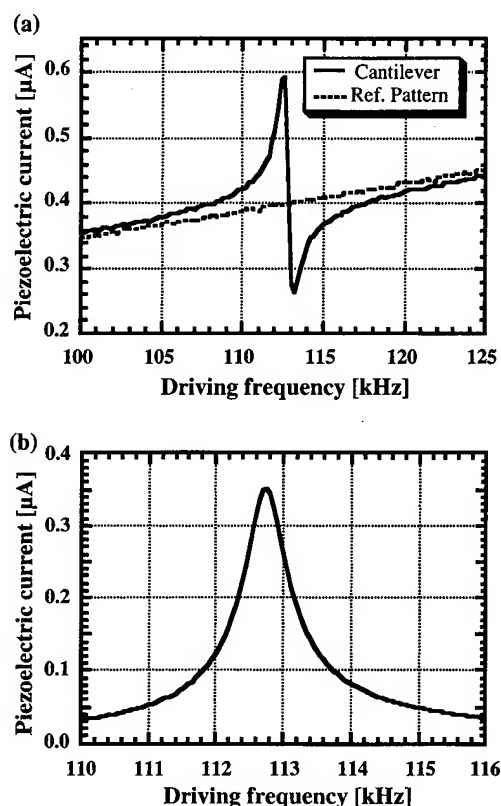


FIG. 2. (a) Current output of a 160- μm -long PZT cantilever and its reference pattern vs the driving frequency. The curves from the cantilever and reference pattern are recorded separately when an ac voltage of 12.5 mV is applied. (b) Piezoelectric current spectrum of a self-excited force sensing 160- μm -long PZT cantilever in air. The excitation ac voltage of 20 mV is directly applied to the cantilever and its reference pattern. The piezoelectric current is measured by subtracting the reference current from the current output of the cantilever via the differential current amplifier. The first mechanical resonance peak is at 112.75 kHz.

by an applied ac driving voltage. The mechanical vibrational amplitude is represented in terms of the piezoelectric current signals from the differential current amplifier based on the mechanism discussed in Sec. II. However the driving voltage can also be applied to a piezoelectric oscillator that is connected with the cantilever; then the cantilever will be excited. Thereafter the piezoelectric charge output will vary with the change of mechanical vibrational amplitude; the difference between charge output and the setup value is taken as the feedback signal to the scanner for z actuation. This is the piezoelectric charge detection method.¹⁴ Using this method the piezoelectric charge output from the cantilever vibrated by the oscillator in liquid can be recorded as the function of driving voltage frequency. The resulting curve is shown in Fig. 4(a). Although we did not immerse the external oscillator into the liquid, the substantial vibration from the bulk silicon base still imposed a strong damping influence on the vibration of the PZT cantilever. The curve is basically similar to the curve demonstrated by Putman *et al.*⁴ Figure 4(b), on the other hand, shows the curve of the piezoelectric current signals versus driving frequency that is measured by the SFM shown in Fig. 3. A broad, but very clear, resonance

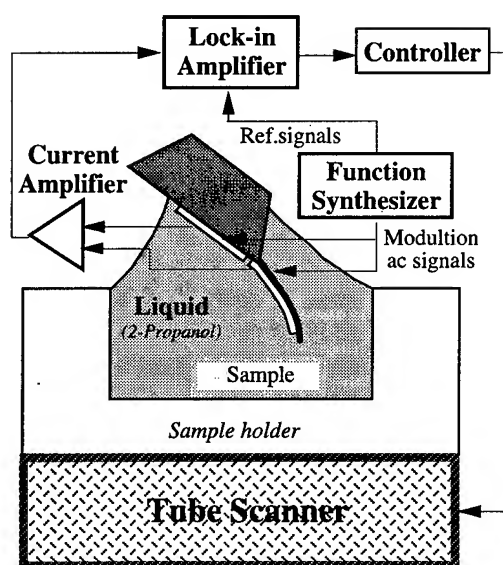


FIG. 3. Schematic diagram of a dynamic SFM using the self-excited force sensing PZT cantilever in liquid. The xy scanning and feedback actuation of tip-sample spacing are executed by the tube scanner. The 2-propanol is used as the liquid in the sample holder for imitating the high viscous environment.

peak near 60 kHz is observed. It indicates that acoustic excitation is almost avoided by using the self-excited cantilever. The incremental trend on the piezoelectric current output as driving frequency is increased is attributed to the fact that the piezoelectric current is proportional to the driving ac voltage frequency. A modified current output versus driving frequency can be calculated from the curve of Fig. 4(b) by deleting the factor ascribed to ac voltage frequency, as shown in Fig. 4(c). The quality factor is then calculated as 6 from Fig. 4(c). Comparing the curve in air, i.e., Fig. 2(b), with the one in liquid, i.e., Fig. 4(b), the earlier resonance frequency drops from 112.7 to 60 kHz, and the quality factor is 200 in air and 6 in liquid, while the driving voltage used in liquid is 7.5 times greater than that in air. The vibrational amplitude versus driving voltage at the resonance can be calculated from the output current based on the method discussed in Ref. 10, in which the vibrational amplitude at the resonance equals the product of the static actuation ability and vibrational quality factor. The corresponding values are 25.6 nm/mV in air and 0.31 nm/mV in liquid.

The damping coefficient per unit length of the cantilever can be obtained from the mass of per unit length, m_e , times the resonance frequency, f_R , divided by the quality factor, Q , i.e., $C \cong m_e f_R / Q$. The damping coefficient of the PZT cantilever in liquid, i.e., 2-propanol, can be calculated as 18 times larger than the one in air. A strong influence by viscous damping in liquid can be concluded.

B. Force curve of the cyclic contact mode and the images derived

Figure 5(a) shows a curve of piezoelectric current signals versus driving frequency by another 160- μm -long cantilever.

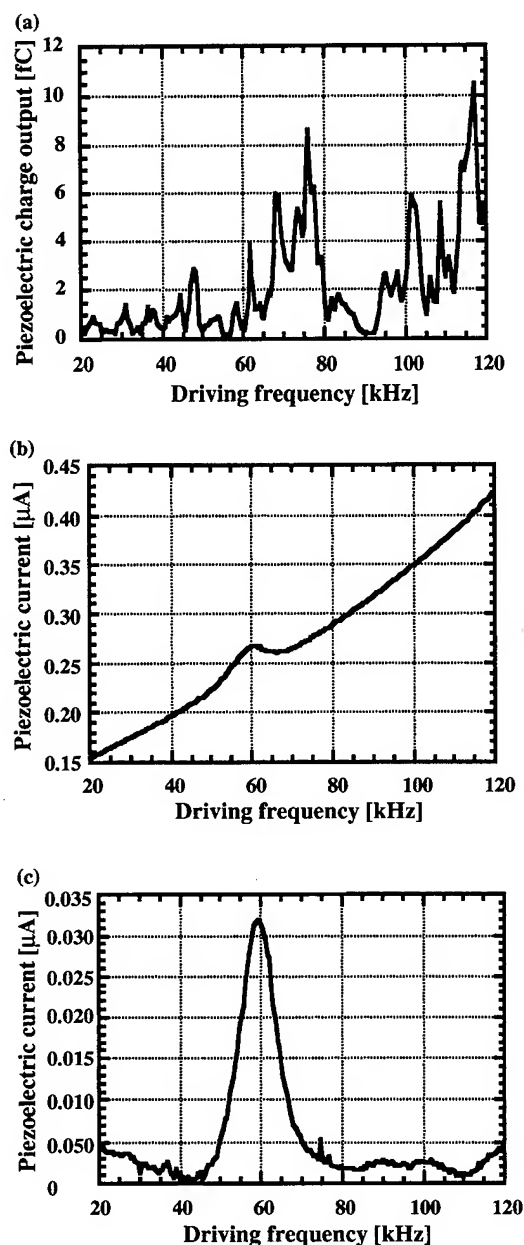


FIG. 4. (a) Piezoelectric charge spectrum of a 160- μm -long PZT cantilever in liquid. The cantilever is excited by an external oscillator. Details of the experimental setup of the piezoelectric charge detection method can be found in Ref. 14. (b) Piezoelectric current spectrum of a self-excited force sensing 160- μm -long PZT cantilever in liquid. The excitation ac voltage is 150 mV. The earlier mechanical resonance peak drops from 112.75 kHz in air [Fig. 2(b)] to 60 kHz in liquid. The cantilever used here is the same one used in (a). (c) A modified spectrum of (b). The quality factor calculated is 6.

A slight shift of resonance frequency is observed between Fig. 5(a) and Fig. 4(b), and the piezoelectric current outputs are somewhat different from the data in Fig. 4(b). These differences may not be due only to the deviation of cantilever characteristics from different cantilevers, but may also be due to the amount of liquid used in experiment setup because the initial amount might be different, and this amount may decrease depending on an increase in the operation time. The force curve derived is shown in Fig. 5(b). It is measured by

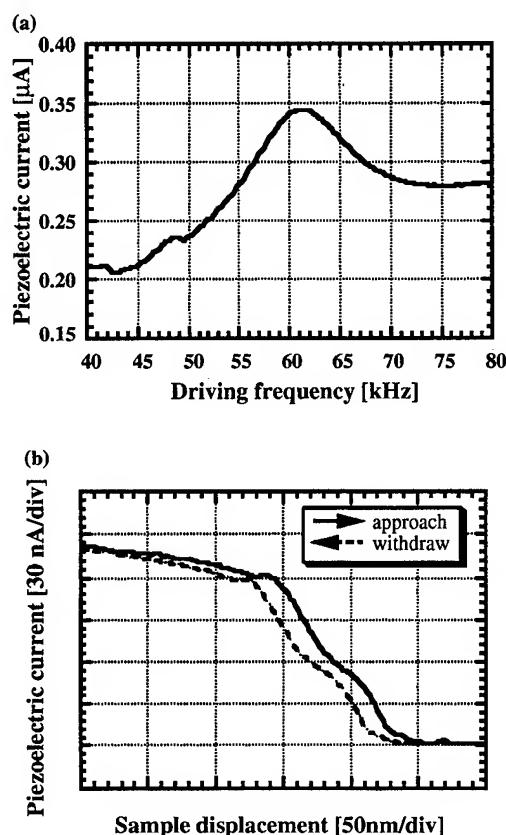
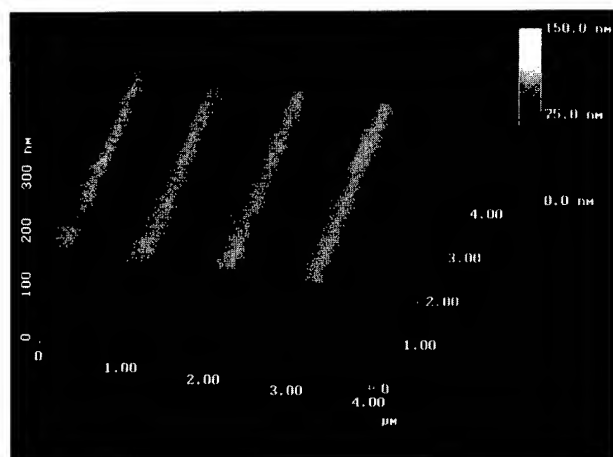


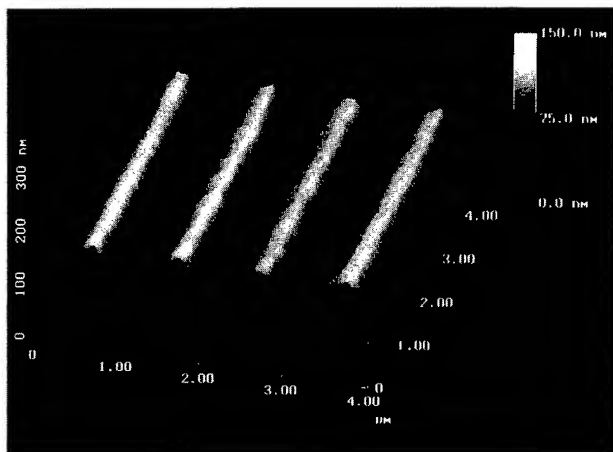
FIG. 5. (a) Piezoelectric current spectrum of a self-excited force sensing 160- μm -long PZT cantilever in liquid with an applied voltage of 375 mV. (b) Force curve trace in liquid of the self-excited PZT cantilever used in (a).

recording the current signals from the current amplifier while the dc actuation voltage is applied to the scanner to control the spacing between the sample surface and the cantilever end. The sensitivity of a 160- μm -long PZT cantilever in liquid can be defined as the slope of the force curve and its value is 1.2 nA/nm. The minimum piezoelectric current output is also derived from the magnified area of a segment in the force curve. Then the amplitude resolution can be estimated by the minimum detectable current output/slope of force curve, i.e., $0.32 \text{ nA} / 1.2 \text{ nA/nm} = 0.27 \text{ nm}$ at a bandwidth of 125 Hz. Because there is an apparatus setup angle, here, 30° , the vertical resolution equals the vertical vector of the amplitude resolution and its value is 2.3 \AA . This vertical resolution is sufficient to image most of the biological samples.

For imaging the sample, a feedback bias is needed to keep the vibrational amplitude at the set point. This bias signal is then recorded as the trace of sample displacement in the z direction, i.e., its topography. Figure 6(a) shows a SFM image of a Au film coated 1.0 μm pitch SiO_2 grating in 2-propanol, taken by cyclic contact SFM with a self-excited force sensing 160- μm -long PZT cantilever. Figure 6(b) shows an image for a similar sample taken by cyclic contact SFM with a self-excited force sensing 125- μm -long PZT cantilever operated in air.¹⁰ In general, these images agree.



(a)



(b)

FIG. 6. (a) A cyclic contact SFM image of a Au film coated 1.0 μm pitch SiO_2 grating in 2-propanol. The scanning rate is 2 Hz and the sampling points are 256×256 . (b) A cyclic contact SFM image of a sample similar to that used in (a) taken in air. The scanning rate is 0.5 Hz and sampling points are 256×256 .

IV. CONCLUSIONS

A new device that includes a PZT reference pattern and a free-standing PZT cantilever is proposed for the realization of a self-excited force sensing PZT cantilever. Using this

structure, the piezoelectric current, which corresponds to the vibration, can be precisely measured. We also demonstrated cyclic contact mode SFM in liquid using this self-excited force sensing PZT cantilever. Since the acoustic excitation from the external oscillator no longer exists, a clear single resonance peak can be observed. A clear grating image in liquid was demonstrated, and it is compatible with the image taken in air. The potential of applying this piezoelectric SFM to the characterization of biological samples in a solution environment is optimistic. Because this piezoelectric SFM of the cyclic contact mode can offer a fast and convenient way for operation in liquid, it may also be possible to reduce the influence of lateral forces and limit deformation on soft biological samples.

ACKNOWLEDGMENT

The authors would like to thank the Japan Science and Technology Corporation (JST) for offering the National Institute postdoctoral fellowship to one of the authors (C.L.).

- ¹G. Binnig, C. F. Quate, and C. H. Gerber, *Phys. Rev. Lett.* **56**, 930 (1986).
- ²M. Radmacher, R. W. Tillmann, M. Fritz, and H. E. Gaub, *Science* **257**, 1900 (1992).
- ³P. K. Hansma, J. P. Cleveland, M. Radmacher, D. A. Walters, P. E. Hillner, M. Bezanilla, M. Fritz, D. Vie, H. G. Hansma, C. B. Prater, J. Massie, L. Fukunaga, J. Gurley, and V. Elings, *Appl. Phys. Lett.* **64**, 1738 (1994).
- ⁴C. A. J. Putman, K. O. Van Der Werf, B. G. De Grooth, N. F. Van Hulst, and J. Greve, *Appl. Phys. Lett.* **64**, 2454 (1994).
- ⁵M. Tortonese, R. C. Barrett, and C. F. Quate, *Appl. Phys. Lett.* **62**, 834 (1993).
- ⁶T. Itoh and T. Suga, *Nanotechnology* **4**, 218 (1993).
- ⁷T. Goddenhenrich, H. Lemke, U. Hartmann, and C. Heiden, *J. Vac. Sci. Technol. A* **8**, 383 (1990).
- ⁸T. Itoh, T. Ohashi, and T. Suga, *IEICE Trans. Electron.* **E78-C**, 146 (1995).
- ⁹T. Itoh and T. Suga, *Sens. Actuators A* **54**, 477 (1996).
- ¹⁰C. Lee, R. Maeda, T. Itoh, and T. Suga, *Proceedings of the 3rd France-Japan/1st Europe-Asia Congress on Mechatronics*, Besancon, France, 1-3 Oct. 1996, Vol. 1, p. 285.
- ¹¹T. Itoh, C. Lee, and T. Suga, *Appl. Phys. Lett.* **69**, 2036 (1996).
- ¹²C. Lee, S. Kawano, T. Itoh, and T. Suga, *J. Mater. Sci.* **31**, 4559 (1996).
- ¹³J.-F. Li, D. Viehland, C. D. E. Lakerman, and D. A. Payne, *J. Mater. Res.* **10**, 1435 (1995).
- ¹⁴C. Lee, T. Itoh, and T. Suga, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **43**, 553 (1996).

Interface imaging by second-harmonic microscopy

M. Flörsheimer,^{a)} M. Bösch,^{b)} Ch. Brillert, M. Wierschem, and H. Fuchs
Physical Institute, University of Münster, D-48149 Münster, Germany

(Received 12 September 1996; accepted 31 March 1997)

Second-harmonic microscopy (SHM) is a fast noncontact surface specific imaging technique for the quantitative *in situ* characterization of a large variety of real interfaces including liquid surfaces, surfaces in vacuum, and buried interfaces. The conventional surface second-harmonic generation experiment reveals the surface order and symmetry averaged over the area illuminated with an intense laser beam. In contrast, the SHM experiment probes the interface laterally resolved. Here we show that the optical nonlinearity of any local feature resolved in an image can be measured quantitatively. The nonlinear optical properties can be interpreted in terms of local surface order and symmetry. The lateral resolution limit of the stigmatic microscope is Abbe's resolution for the generated SH light with a wavelength of, e.g., 532 nm. The contrast mechanism can also be applied to a scanning optical near field microscope. © 1997 American Vacuum Society. [S0734-211X(97)04304-7]

I. INTRODUCTION

Second-order nonlinear optical techniques are efficient remote sensing methods for the *in situ* investigation of a large variety of real interfaces including liquid surfaces, surfaces in vacuum, and buried interfaces.¹⁻⁷ In the bulk of centrosymmetric or isotropic media, such as silicon, glass, or liquids, the second-order optical processes are symmetry forbidden if the electric-dipole interaction between light and matter is the predominant mechanism. Hence, the nonlinear optical signal emerges selectively from the surfaces or interfaces where the bulk symmetry is necessarily broken.¹

The nonlinear processes require high intensities of the probing light. Such intensities can be provided by readily available pulsed laser sources. The simplest experiment is second-harmonic generation (SHG) where a single laser beam is used to produce frequency-doubled light at the interface. The generated signal carries information on the symmetry and order of the interface.

In the conventional surface SHG experiment, the frequency-doubled light is measured integrally from the illuminated spot.¹⁻⁷ Hence, the average surface order and symmetry within the spot can be determined. SH imaging, however, reveals the lateral structure of the interface with far field optical resolution.⁸ The range and changes of orientational order can be directly observed. The crystallographic structure of any local feature resolved in an image can be determined. If the optical nonlinearity of the interface is sufficiently high, an image can be created applying a single laser pulse with a time duration of, e.g., 100 ps.⁹ Thus, SH microscopy is a promising method for the fast, quantitative, and remote sensing investigation of surfaces, interfaces, and ultrathin films.

II. THEORY

The SH intensity $I^{(2\omega)}$ generated in a surface analytical experiment is given by¹⁰

$$I^{(2\omega)} = \frac{2 \omega^2 L^2}{\epsilon_0 c^3 \cos^2 \Omega} |d_{\text{eff}}|^2 (I^{(\omega)})^2 \quad (1)$$

with d_{eff} the effective susceptibility, $I^{(\omega)}$ the intensity of the fundamental beam (frequency ω), Ω the angle of incidence, L the thickness of the SH active layer, ϵ_0 the vacuum permittivity, and c the speed of light in vacuum. The effective susceptibility d_{eff} is defined as

$$d_{\text{eff}} = \mathbf{e}^{(2\omega)} \vec{d} : \mathbf{e}^{(\omega)} \mathbf{e}^{(\omega)} = \sum_{I,J,K} e_I^{(2\omega)} d_{IJK} e_J^{(\omega)} e_K^{(\omega)}. \quad (2)$$

The tensor \vec{d} reflects the actual macroscopic material properties. Its elements are given by d_{IJK} where the index I refers to the coordinates of the SH field. The quantities $\mathbf{e}^{(2\omega)}$ and $\mathbf{e}^{(\omega)}$ are Fresnel factors of the interface multiplied by unit vectors which describe the orientation of the electric fields of the SH and the fundamental light, respectively. The effective susceptibility is a scalar which depends on the geometry used in an experiment, and on the tensor elements which are relevant for this special geometry. \vec{d} is a third-rank tensor with a maximum of 18 independent elements.¹¹ The corresponding material property in linear optics is given by the dielectric tensor.

According to Eq. (1), the susceptibility \vec{d} cannot be determined directly in an experiment. The quantity which can be measured is the SH intensity radiated from the illuminated specimen. From the intensity, the effective susceptibility can be calculated. Then, the experimental geometry can be altered such that an additional effective susceptibility can be determined. From N known effective susceptibilities, N ele-

^{a)}Electronic mail: Florschl@NWZ.Uni-Muenster.De

^{b)}Present address: Swiss Federal Institute of Technology, ETH-Hönggerberg, CH-8093, Zürich.

ments of the tensor \vec{d} can be calculated. Finally, the macroscopic property \vec{d} can be interpreted on the molecular or atomic level.

The microscopic property corresponding to the susceptibility as described by \vec{d} is the hyperpolarizability $\vec{\beta}$. In the simplest case, a molecule possesses a single axis z of high hyperpolarizability β_z .¹² If molecular interactions can be neglected, the macroscopic nonlinearities can be calculated from the molecular nonlinearity by a simple projection of the molecular onto the macroscopic coordinates and by summation over all molecules.¹² We have

$$d_{IJK} = \frac{1}{V_{\text{molecules}}} \sum \langle \hat{I} \cdot \hat{z} \rangle \langle \hat{J} \cdot \hat{z} \rangle \langle \hat{K} \cdot \hat{z} \rangle \beta_z \quad (3)$$

with \hat{I} , \hat{J} , \hat{K} , and \hat{z} the unit vectors of the coordinates, and V the volume of the film. The angle brackets refer to an orientational average over the molecules. In linear optics, there is a similar relation between the dielectric tensor and the molecular polarizabilities. Generally however, more information on the molecular order can be obtained in nonlinear optics because the higher order susceptibilities possess more independent elements (up to 18 for SHG compared to three elements of the diagonalized dielectric tensor in linear optics).

Summarizing, the relevant quantities of a SHG experiment are determined in the following sequence:

$$I^{(2\omega)} \rightarrow d_{\text{eff}} \xrightarrow{\vec{d}} \vec{\beta}.$$

In the conventional SHG experiment, all these quantities are averaged properties describing the entire interface or a large section. In the SHM experiment, we treat any local feature such as a domain as a sample whose susceptibility tensor can be determined and interpreted.

III. CONVENTIONAL AND MICROSCOPY EXPERIMENT

Various geometries have been used in conventional (integrating) SH generation experiments in order to determine the full susceptibility tensor of a sample. Usually a spot of the specimen is illuminated with a laser beam of constant pulse intensity (Fig. 1). A defined fundamental polarization is selected [s or p in Figs. 1(a) and 1(b)]. The generated SH signal passes an analyzer which also selects a defined polarization. The intensity of the SH light in this state is measured. The geometry can be modified by the rotation of the sample about one of its axes [Figs. 1(a) and 1(b)] such that the effective susceptibility can be determined for various sample orientations. In the example of a liquid surface, for example, [Fig. 1(c)] it is difficult to rotate the specimen. In such cases, for instance, the fundamental polarization can be rotated.

From all these geometries of conventional experiments, the set up of a SHM experiment can be derived. Here, we use a transmitted light geometry (Fig. 2). Scanning of the sample with a tightly focused laser beam is not required. The entire

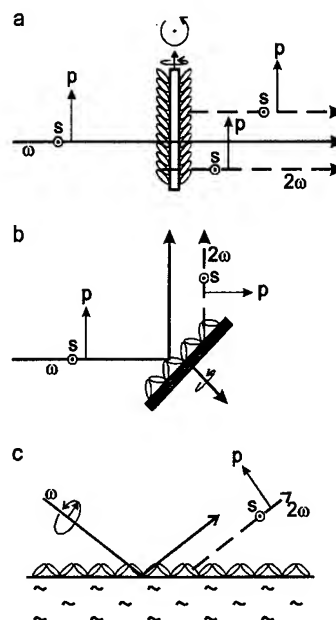


FIG. 1. Conventional surface second-harmonic generation experiments, transmitted (a) and reflected light geometries (b) and (c).

area of interest with a typical diameter of $400 \mu\text{m}$ is illuminated employing a pulsed laser (Spectra Physics GCR 190, wavelength 1064 nm , pulse duration 8 ns , repetition rate 30 Hz , intensity $\sim 200 \text{ MW cm}^{-2}$). The SH light is used to create a magnified image of the interface structure on the target of an intensified CCD camera (Spectroscopy Instruments ICCD-576 G/RB). The two-dimensional detector measures the SH photons with 2.2×10^5 channels in parallel. The response of the camera is calibrated.¹³

Applying normal incidence of the fundamental light the probe is sensitive to the noncentrosymmetry of the monolayer in the surface plane since the fundamental light is polarized in this plane. Below, the performance of our SHM will be demonstrated with the example of a Langmuir mono-

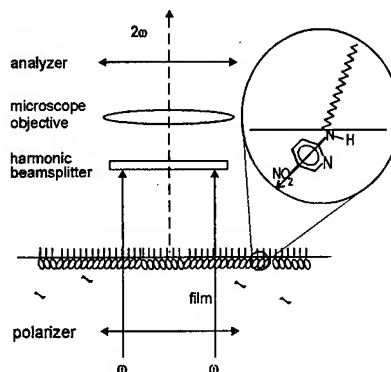


FIG. 2. Scheme of second-harmonic microscope. Here the order of a Langmuir monolayer at a water surface is probed. The axis of the highest molecular hyperpolarizability (charge transfer axis) is indicated with an arrow in the inset.

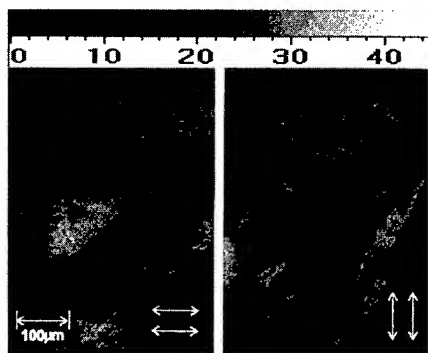


FIG. 3. Second-harmonic micrographs of a Langmuir monolayer of 2-docosylamino-5-nitropyridine (DCANP) at a water surface. The grey scale is calibrated in units of the effective susceptibility d_{eff} of the monofilm (pm V^{-1}). The polarizer orientations are indicated by the arrows.

layer of 2-docosylamino-5-nitropyridine (DCANP) at a water surface. The chemical formula is given in Fig. 2. The signal from the monofilm proved to be much more intense than the signal from the bare water surface. Thus, the latter can be neglected.

IV. RESULTS AND INTERPRETATION

SH micrographs of a DCANP monolayer are shown in Fig. 3. The monolayer was observed between parallel polarizers whose orientation is indicated. An image was taken within 10 s (300 laser pulses). The grey scale is calibrated in units of the effective susceptibility (pm V^{-1}) of the film. Remember that the generated SH intensity $I^{(2\omega)}$ is proportional to $|d_{\text{eff}}|^2$ as given by Eq. (1). The textures seen in the images can be described as irregular, elongated, dendritic structures. Contrast inversion occurs if the polarizers are rotated by 90° . Radiation damage due to the high light intensities was not observed.

The contrast can originate from different phenomena. In the most general case, the molecular order and arrangement varies laterally. Another possibility is that all the local structural features resolved in the images possess the same or a similar crystallographic symmetry and order. This means that the susceptibility tensor which describes the local features is identical for all the features. In such an example, the lateral variation of the effective susceptibility originates from orientational variations of the crystallographic axes. This interpretation is valid for all polycrystalline surfaces and for various liquid crystalline mesophases.

We assume this interpretation to be also valid in the example of the DCANP monolayers. In addition, we see that all the features resolved must be noncentrosymmetric in the surface plane, that is, a polar order exists in the plane of the film. As a sample, we can define any local feature with, for example, a homogeneous effective susceptibility. We assume a single polar axis corresponding to a single susceptibility tensor element is sufficient to describe the in-plane order of all these local features. An example of such an object with its polar axis Z and the susceptibility d_{zzz} is depicted in Fig. 4. The polarizer orientations are given in a laboratory frame

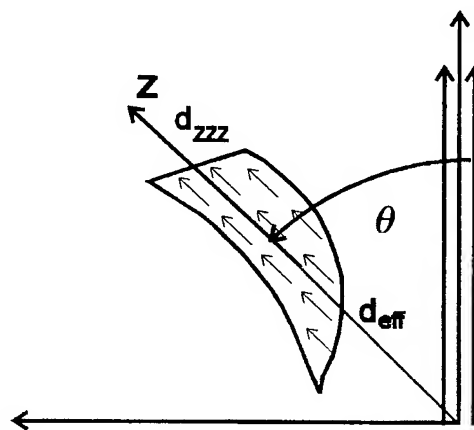


FIG. 4. Example of a local feature with its macroscopic polar axis Z and its orientation θ relative to the polarizers. The nonlinear optical data are interpreted such that the charge transfer axes of the molecules in the local feature (small arrows) are oriented parallel to the macroscopic polar axis. A laboratory coordinate system lying in the surface plane is also given. The contrast between different local features in an image originates from the lateral variation of the molecular orientation relative to the laboratory frame.

lying in the surface plane. Using this model, the features whose polar axes are oriented favorably with respect to the polarizers should appear bright in an image. If all these assumptions hold, a contrast inversion is expected for a rotation of the polarizers by 90° . This agrees well with the observation indicating that the model is reasonable.

The interpretation can also be verified quantitatively. If a single susceptibility tensor element d_{zzz} is sufficient to describe a local feature, Eq. (2) can be written as

$$d_{\text{eff}} = F^{(2\omega)} (F^{(\omega)})^2 \cos^3 \theta d_{zzz}, \quad (4)$$

where $F^{(2\omega)}$ and $F^{(\omega)}$ are the Fresnel factors for the interface.¹⁰ θ is the angle between the polarizer orientation and the Z axis. The cosine factors account for the projection of the two fundamental fields from the polarizer orientation onto the Z axis of a local feature and for the projection of the generated SH field from the polar axis onto the analyzer of the microscope. The validity of Eq. (4) can be proven by synchronous rotation of the polarizers in an experiment and measuring of d_{eff} in dependence on θ . These experiments will be described in detail elsewhere.¹³ The result obtained from several experiments with different DCANP Langmuir films is that Eq. (4) fits the data well. We obtain the following results:

- (1) The in-plane structure of all the local features can be described with a single predominant susceptibility d_{zzz} ,
- (2) the value of this susceptibility is identical for all the features, and
- (3) its magnitude¹⁴ is $|d_{zzz}| = 40 \pm 4 \text{ pm V}^{-1}$.

In the following, the measured macroscopic nonlinearities will be interpreted molecularly. The hyperpolarizability of a molecule like DCANP is due to an intramolecular charge-transfer (CT) group:¹² a π -electron system (pyridine ring) that is substituted by an electron donor (amino group) and an

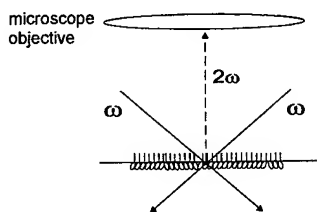


FIG. 5. Second-harmonic microscope taking advantage of a noncollinear geometry.

electron acceptor (nitro group). The axis of highest hyperpolarizability corresponds to the CT axis. This molecular axis is indicated in Fig. 2 by an arrow. Since there are both a single molecular and a single macroscopic nonlinearity, the CT axes of all the molecules in a local feature must be oriented parallel to macroscopic polar axis of this feature. This means that the projection factors in Eq. (3) amount to one. The local molecular order is indicated in Fig. 4 using the small arrows.

The contrast of the SH micrographs can then be explained as follows: The orientation of the molecules (or the brightness of the images) is maintained over 10 μm dimensions. Over longer distances, the orientational order changes either abruptly or continuously. Using Eq. (4), the measured two-dimensional distribution of the effective susceptibility can be readily transformed into the corresponding molecular orientation field.

V. DISCUSSION

The parallel orientation of the molecular π -electron system agrees well with electron diffraction data obtained from DCANP films which were transferred onto solid substrates.¹⁵ Advantages of the optical technique are: No sample preparation is required, interface selectivity is inherently provided, and lateral resolution can easily be obtained. SHM is not restricted to the characterization of Langmuir films. It can be applied to all kinds of interfaces accessible to light. SHM is also not restricted to simple cases such as DCANP monolayers where a single susceptibility is sufficient to describe the in-plane structure. In the case of the lowest possible symmetry of an acentric feature (C_1), there are up to six independent susceptibilities required to describe the in-plane noncentrosymmetry. It can be shown that all these susceptibilities can be measured.¹⁶ At an interface, however, it is highly desirable to be also sensitive to the polar order in normal direction.

Normal dipoles can be probed applying an oblique incidence of the fundamental light. Then, however, the SH light is radiated also under an oblique angle collinearly with the reflected fundamental light. [See Figs. 1(b) and 1(c) for the conventional experiment.] Hence, the objective axis of a stigmatic microscope must also be tilted with respect to the surface. As a result, a highly magnified image with high resolution can only be obtained from a small stripe in the object field. Recently, we have, however, realized a noncollinear geometry where two fundamental beams are used¹⁷

(Fig. 5). Due to the conservation of momentum, the SH photons are radiated normal to the surface. This experiment combines a number of advantages:

- (1) Additional susceptibilities can be determined quantitatively due to the nonnormal incidence of the fundamental light.
- (2) Abbe's resolution for the SH light can be obtained in the entire object field due to the objective axis normal to the surface.
- (3) Specimen can be probed in reflection. Hence, nontransparent interfaces such as metal and semiconductor surfaces can be measured.

In order to obtain higher lateral resolution, the contrast mechanism could, in principle, also be applied to a scanning optical near field microscope. Very high laser intensities, however, are required in order to obtain a sufficient number of SH photons from the small areas observed. For an estimation of the SH signals we assume the detection of the SH light through an aperture with a diameter of 50 nm. The number of SH photons generated within a surface section corresponding to this aperture can be calculated using Eq.

(1). For the illumination of the surface a freely propagating laser beam may be utilized. Also, a total internal reflection geometry could be applied. For a fundamental intensity of 100 TW cm^{-1} , a pulse duration of 400 fs, a repetition rate of 100 MHz, a wavelength of 1064 nm, a thickness of the SH active film of 2 nm, an effective susceptibility of 1 pm V^{-1} , an integration time of 10^{-3} s, and for normal incidence of the laser beam we obtain a number of 2.2×10^7 SH photons from the area adjacent to the aperture. This number is sufficient for imaging. The high fundamental intensities required can be applied especially to biological samples if fs laser pulses are used.

In summary, we have shown that the lateral structure of an interface, a surface, or an ultrathin film can be imaged by SHM. The susceptibility of all the local features resolved in the images can be measured quantitatively. The two-dimensional distribution of the susceptibilities can be readily transformed into the corresponding molecular orientation field. The technique can be applied to all kinds of real interfaces accessible to light.

ACKNOWLEDGMENT

This work was supported by the Federal Ministry for Science and Research of Northrhine Westfalia and from the Volkswagen Foundation.

¹Y. R. Shen, *Surf. Sci.* **299/300**, 551 (1994).

²R. B. Hall, J. N. Russell, J. Miragliotta, and P. R. Rabinowitz, in *Chemistry and Physics of Solid Surfaces*, edited by R. Vanselow and R. Howe (Springer, Berlin, 1990), Vol. 8, pp. 87–132.

³V. Vogel and Y. R. Shen, *Annu. Rev. Mater. Sci.* **21**, 515 (1991).

⁴M. Buck, *Appl. Phys. A* **55**, 395 (1992).

⁵G. Lüpke and G. Marowski, *Phys. Bl.* **49**, 285 (1993).

⁶R. M. Corn and D. A. Higgins, *Chem. Rev.* **94**, 107 (1994).

⁷J. F. McGilp, in *Progress in Surface Science*, edited by S. G. Davison *et al.* (Elsevier Science, Oxford, 1995), Vol. 49, pp. 1–106.

- ⁸M. Flörsheimer, C. Radüge, H. Salmen, M. Bösch, R. Terbrack, and H. Fuchs, *Thin Solid Films* **284–285**, 659 (1996).
- ⁹M. Flörsheimer, H. Looser, M. Kúpfer, and P. Günter, *Thin Solid Films* **244**, 1001 (1994).
- ¹⁰V. Mizrahi and J. E. Sipe, *J. Opt. Soc. Am. B* **5**, 660 (1988).
- ¹¹R. W. Boyd, *Nonlinear Optics* (Academic, Boston, 1992).
- ¹²Ch. Bosshard, K. Sutter, Ph. Prêtre, J. Hulliger, M. Flörsheimer, P. Kaatz, and P. Günter, *Organic Nonlinear Optical Materials*, *Advances in Nonlinear Optics*, Vol. 1, edited by A. F. Garito and F. Kajzar (Gordon and Breach, Basel, 1995).
- ¹³M. Flörsheimer, M. Bösch, Ch. Brillert, M. Wierschem, and H. Fuchs, *Supramolecular Science* (in press).
- ¹⁴For comparison: The highest susceptibility tensor element of LiNbO_3 amounts to 34 pm V^{-1} .
- ¹⁵F. Klinkhammer, Ph.D. thesis, Mainz, 1995; I. R. Peterson (personal communication).
- ¹⁶M. Flörsheimer, M. Bösch, Ch. Brillert, M. Wierschem, H. Salmen, and H. Fuchs (unpublished).
- ¹⁷M. Flörsheimer, M. Bösch, and H. Fuchs, *Adv. Mater.* (submitted).

Scanning acoustic tunneling microscopy and spectroscopy: A probing tool for acoustic surface oscillations

T. Hesjedal,^{a)} E. Chilla, and H.-J. Fröhlich
Paul-Drude-Institute for Solid-State Electronics, D-10117 Berlin, Germany

(Received 8 September 1996; accepted 21 April 1997)

A method is presented for the universal probing of surface acoustic waves (SAWs). For measuring high frequency SAWs by scanning tunneling microscopy (STM) a stroboscopic snapshot technique was employed, named scanning acoustic tunneling microscopy. The amplitude and the phase of a state of oscillation within a SAW field are a superposition of the surface topography and the oscillation trajectory. Measuring with atomic resolution the observed contrast can be understood by a spherical tunneling model. A STM based system is proposed that reaches submicron resolution for the quantitative evaluation of elastic constants. With this system the velocity dispersion is obtained from the detection of laser generated SAW pulses by a broadband STM. Scanning acoustic tunneling spectroscopy using different acoustic modes opens a door for quantitative studies of nanoscale structures. © 1997 American Vacuum Society. [S0734-211X(97)10004-X]

I. INTRODUCTION

With miniaturization in electronics a need for mechanically characterizing the thin films used arose.¹ These films often limit device durability due to their high inner stresses. A well-suited method for the investigation of thin films is based on surface acoustic waves (SAWs) since they are confined within some wavelengths in the vicinity of the surface. Although high frequency SAWs (typically MHz to GHz) are nondispersive when propagating on an elastic half-space, a layer attached to the bulk substrate causes dispersion. The dispersion curve is an implicit function of the elastic constants of substrate and layers. Hence the measurement of the frequency dependent phase velocity over a wide frequency range, also called SAW spectroscopy,² offers a unique tool for studying the elastic properties of thin films.

In general, any combination of an acoustic source with an acoustic receiver provides a setup for SAW spectroscopy. The most desired arrangements are those based on noncontact broadband transducers delivering the dispersion curve out of a single acoustic pulse. But most SAW transducer principles are restricted to piezoelectric substrates and to certain crystal cuts, or are limited in their bandwidth. The most important techniques for SAW generation are laser optical excitation and interdigital transducers (IDTs).

Laser optical excitation allows the generation of either broadband acoustic pulses by pulsed lasers or narrow-bandwidth waves by scanning interference fringes (SIFs)³ on almost arbitrary substrates. The combination of pulsed laser excitation with laser probing, piezoelectric foil detectors or IDT detectors are standard setups for nondestructive evaluation purposes.⁴ Yet these spectroscopy techniques are not supposed to attain lateral resolution. Even standard 225 MHz line-focus acoustic microscopy, capable of measuring SAW velocity, is not heading towards the submicron scale.⁵

With the invention of molecular beam epitaxy the generation of vertically as well as of laterally confined structures

was accomplished.⁶ For the study of such very thin layers high acoustic frequencies in the super-GHz range are necessary, and this is offered by ultrafast optical pump-and-probe methods involving acoustic bulk waves of up to 100 GHz.⁷

For elastically characterizing laterally confined structures we propose a SAW dispersion measurement method based on the reduction of the size of the excitation source and/or the detecting probe by means of scanning probe microscopy. The most sophisticated system stage will combine pulsed laser excitation with transient scanning tunneling microscopy (STM) methods for both conducting and nonconducting surfaces. For two reasons we started with IDT excited SAWs. First, IDT excitation guarantees well-defined surface oscillation conditions enabling us to model the influence of topography and acoustic modes on the microscope's contrast. Second, the excitation of continuous waves which was recently optically realized by SIF allows one to circumvent some of the key problems in measuring fast phenomena with bandwidth limited controller electronics. Using the nonlinear force and tunneling curves as frequency demodulators or mixers in the same way as a crystal radio detector, scanning acoustic force microscopy (SAFM) and scanning acoustic tunneling microscopy (SATM), respectively, transfer the rf wave's information to easy-handling kHz frequencies.⁸

In this article heterodyne detection as well as the discussion of transient STM methods will be presented. Then the influence of topographic features on the SATM contrast will be discussed, both theoretically and experimentally. Atomic resolution measurements show a snapshot of the atomic oscillation trajectories. In addition a system is proposed combining laser optical pulse excitation with a transient STM; this is called scanning acoustic tunneling spectroscopy (SATS).

II. SATM: DETECTION OF ACOUSTIC WAVES

The nonlinearity describing the tunneling of electrons from a sharp tip to a sample surface can be utilized as an

^{a)}Electronic mail: hesjedal@pdi.wias-berlin.de

effective high-frequency signal mixer. When a propagating SAW with the normal component u_z at ω_{SAW} passes the tip modulating the undisturbed tip-to-sample distance d_0 in the form $d = d_0 + u_z \sin(\omega_{\text{SAW}} t + \varphi)$ the applying of a slightly frequency detuned voltage $V_1 \sin(\omega_{\text{clock}} t)$ at ω_{clock} additionally to the gap voltage V_0 leads to the following tunneling current at the lowest difference frequency $\Delta\omega = |\omega_{\text{SAW}} - \omega_{\text{clock}}|$:

$$I \sim \frac{u_z}{d_0} \frac{V_1}{V_0} (\text{const} \cdot d_0 + 1) \cos(\Delta\omega t + \varphi_{\text{SAW}}). \quad (1)$$

The local SAW phase φ_{SAW} is linearly dependent on the propagation length. By analyzing the tunneling current with a lock-in amplifier at $\Delta\omega$ the obtained images can be seen as a stroboscopic snapshot of the oscillating surface.

Some of the main advantages of such a heterodyning technique are the avoidance of transducer-tip crosstalk influences and strong displacement currents between the tip and the metallized surface by measuring at the difference frequency.⁹

The need for ultrafast STM microscopes comes generally from the fact that small structures exhibit short time scales. For SAW detection in the frequency range from MHz to some GHz at least nanosecond time resolution should be reached. Although the inherent tunneling time was estimated to be 10 fs or lower, the electromagnetic feedthrough and stray capacities have to be overcome. Adapting ultrafast optical pump-and-probe techniques a fast gate signal is needed to lead through the nonlinearity of the tunneling curve to a time averaged and hence conventionally measurable signal. For gating several approaches were chosen: magnetostrictive mechanical gating,¹⁰ optical gating through a photoconductive switch at the tip¹¹ and voltage gating through a photoconductive switch at the sample. As SAW excitation by pulsed lasers offers a fast trigger signal, photoconductive switch techniques can be implemented.

A different approach is given through the direct high-frequency acoustic pulse probing by means of an ac STM.¹² Possibly through the charge transfer from and to the sample the imaging of insulating surfaces was achieved. For broadband laser and IDT SAW spectroscopy we constructed a Beetle-style¹³ ac STM for a frequency range from 0 to 3 GHz.¹⁴

III. SATM: INFLUENCE OF TOPOGRAPHY

Since the atomic oscillation of a SAW is polarized and the STM tip is seeing a snapshot of that oscillation, the direction of the tunneling with respect to the normal component of surface oscillation is of great importance. Hence, the surface topography, as well as the shape of the atomic oscillation, influence amplitude and phase contrast. Assuming the tip to be a point source tunneling on corrugated surfaces can easily be written in spherical coordinates $I(\vec{d}) = I(d_0, \Phi, \Theta)$. Introducing the wave's displacement components u_i in spherical coordinates too, for $u_i \ll d_0$ the tunneling current density can be written in a linear approximation as

$$\vec{j} \sim \vec{u} \begin{pmatrix} \cos \Phi \sin \Theta \\ \sin \Phi \sin \Theta \\ \cos \Theta \end{pmatrix} \sin \left[\Delta\omega t + \begin{pmatrix} \varphi_x \\ \varphi_y \\ 0 \end{pmatrix} + \varphi_{\text{SAW}} \right], \quad (2)$$

with φ_x, φ_y the phase differences of u_x and u_y with respect to u_z .

SAWs are waves that propagate along the surface of an elastic solid and that decay in depth within some wavelengths. The various types of SAWs differ in their polarization, phase velocity, dispersion and decay behavior. As seen from Equation (2) any kind of differently polarized SAWs, i.e., particle displacements in all three directions, are in general measurable by SATM. This makes the method unique compared with standard optical which is unable to detect in-plane oscillation components.

For calculating the possible acoustic modes in a system consisting of a substrate and a certain number of layers a set of mechanical and electrical boundary conditions has to be solved. A software code based on Adler's matrix method was developed to calculate multilayer systems.¹⁵

In the special case of Rayleigh waves on the Y - Z crystal plane of LiNbO_3 the elliptical oscillation is restricted to the x - z plane with x being the SAW propagation direction, i.e., $u_y = 0$. The phase difference between u_x and u_z is $\pi/2$. The amplitude A and the phase Ψ of the tunneling signal at the difference frequency can consequently be written as⁹

$$\Psi = \arctan \left(\frac{u_x}{u_z} \cos \Phi \tan \Theta \right) + \varphi_{\text{SAW}}, \quad (3)$$

$$A = \sqrt{u_x^2 \cos^2 \Phi \sin^2 \Theta + u_z^2 \cos^2 \Theta}. \quad (4)$$

For tunneling vectors parallel to the overall surface normal the topography related phase tends to 0 and only the amplitude and the phase of the u_z component of the SAW are measured. The maximum contrast of amplitude and phase occur in the case of topographical gradients oriented perpendicular and parallel to the propagation direction, respectively.

When measuring the topography of a smooth surface with atomic resolution, a wave propagating along this surface influences the contrast. For wave frequencies much higher than the bandwidth of the STM the topography is an average over many oscillation cycles. For longitudinal bulk waves the atomic contrast simple smears out depending on the wave's amplitude.¹⁶ As the particle displacement is elliptically polarized for SAWs on thin Au layers on YZ - LiNbO_3 , we found a smearing out of the contrast along the wave's propagation direction as indicated by the arrow in Fig. 1.

IV. SATM: ATOMIC RESOLUTION

The following measurements were performed on ≈ 100 nm thick thermally evaporated Au layers on YZ - LiNbO_3 . The SAWs with $\lambda = 88 \mu\text{m}$ were excited using IDTs at 39.50 MHz. A 10 dBm driving signal was fed without electrical matching to the IDT generating sub-Å surface oscillation

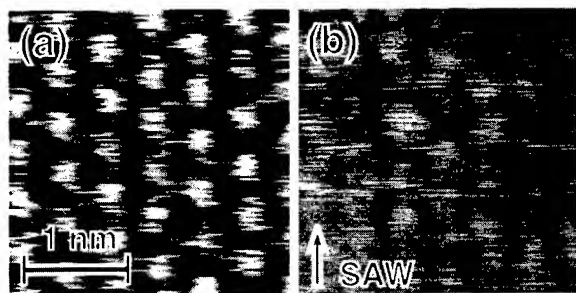


FIG. 1. Undisturbed topography (a) of a Au(111) surface. An elliptically polarized particle displacement due to a propagating SAW (wave vector indicated by the arrow) leads to a directed smearing out of the atomic contrast (b).

amplitudes. The electrical clock frequency was chosen to 39.51 MHz at 1 V rms leading to the signal of interest at the lowest mixing frequency at 10 kHz.

Figure 2(a) shows a $80 \times 80 \text{ nm}^2$ constant current image of the topography. The gold grains have typical dimensions of some tens of nanometers and are oriented almost parallel to each other which may be an effect of the piezoelectricity of the substrate. The simultaneously measured phase image [Fig. 2(b)] first reproduces the grain boundaries with their gradients parallel to the wave's propagation direction (arrow). Second, a fine structure, visible on the individual grains, reveals almost periodic structural features with a distance of 2–5 nm between them. In several areas steps which were identified as being of atomic scale in height (not necessarily monoatomic steps) are visible. The flatter areas especially those like in the middle of the images exhibit larger step-free terraces that cannot be seen in the topographical image. To some extent this measurement mode is similar to scanner dithering differential STM which is also being sensitive on small-scale but steep transitions.¹⁷ Differential-type STMs further improve the signal-to-noise ratio by transferring the relevant signal contributions to a higher frequency band that has lower noise contributions.

The smallest surface feature to be studied is the single atom. The atomically flat Au surfaces needed for this type of measurements were prepared by annealing the sample in an oxygen atmosphere. The resulting (111) facets were a few 100 nm wide.

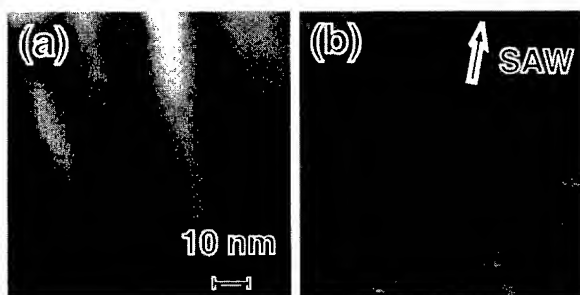


FIG. 2. SATM topography (a) and phase (b) image of a Au surface. The phase image clearly reveals a stepped structure even on Au grains. The SAW propagation direction is indicated by the arrow.

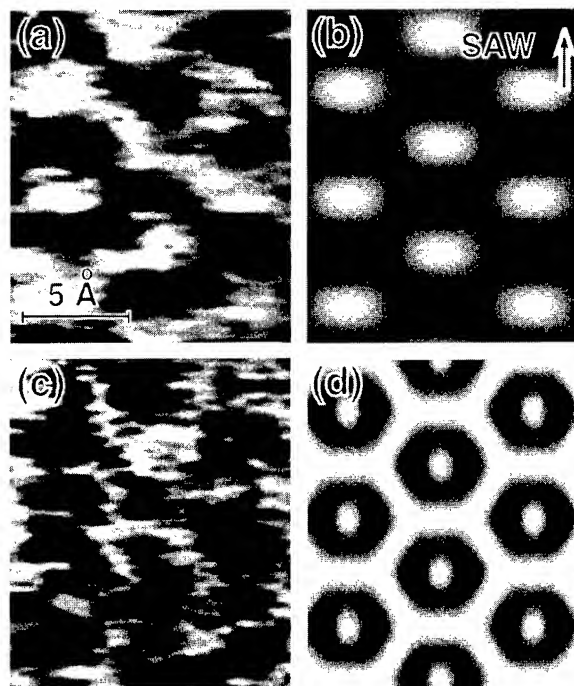


FIG. 3. Comparison of the measured and slightly smoothed (a) and (c) with the modeled (b) and (d) phase and amplitude, respectively, revealing the hexagonal lattice structure. The wave's propagation direction is indicated by an arrow (b).

Although the atomic topography contrast vanishes (Fig. 1), the phase and the amplitude signal [Figs. 3(a) and 3(c)] show defined structures at the positions of the atoms in the hexagonal lattice. A slight smoothing was applied to the data. In the case of phase the atomic contrast structure switches from bright to dark at the atom's equilibrium position in according to the direction of the propagating wave (arrow). Due to adsorbate contamination the amplitude image is much more distorted. Nevertheless ring-type atomic structures are visible. The transferring of the model as expressed in Equations (3) and (4) to the atomic scale was performed by assuming the atoms to be Gaussian-shaped electron distributions around the hexagonal lattice positions. For a calculated ratio of the axes of the oscillation ellipse of $u_z:u_x=0.6:1$ the modeled phase and amplitude images are drawn in Figs. 3(b) and 3(d). Now the phase contrast can be understood in the following way: when scanning over an atom in the propagation direction, the phase reaches a maximum for the displacement vector seen by the snapshot and tunneling vector being parallel; in front and behind the atom the phase is leading and is retarded, respectively. Contrary to that the amplitude is reproducing the reciprocal lattice around the atom's position with an additional maximum in the middle. These aspects were not clearly resolved by the experiment. Nevertheless ultrahigh vacuum (UHV) measurements should reduce the amplitude noise. Since the shape of the oscillation trajectory is reflecting the elastic constants of the probed volume too, phase and amplitude measurements on the atomic scale are also providing a path to elastic microscopy.¹⁸

V. SATS: SCANNING ACOUSTIC TUNNELING SPECTROSCOPY

The combination of high-resolution wave probing SATM with thin film analysis SAW spectroscopy provides a base for future nanometer scale quantitative elastic characterization.

The excitation of SAWs can be performed in several ways, e.g., IDT chirps, vibrating miniaturized objects in contact with the surface like a STM tip,¹⁹ direct piezoelectric coupling, bulk wave mode conversion and optical methods. The applicability of the system is mainly determined by the need of exciting SAWs on any substrate material. Therefore, photoacoustics offers well adapted methods of contact-free SAW excitation. Then two possible approaches to SATS should be discussed: the combination of continuous wave SIFs with heterodyne SATM and the combination of pulsed laser SAW excitation with broadband STMs. As SIFs are not easy to handle, our future work will be dedicated to broadband STMs.

In our first experimental setup we combined a nitrogen pulse laser with a Beetle-type broadband STM. According to the relation $\tau \cdot \Delta f \approx 1$ the 2.5 ns pulse laser chosen should be able to excite 400 MHz bandwidth acoustic pulses.²⁰ Since broadband STMs were demonstrated to work with the 20 GHz bandwidth,¹² the practical limit of such a setup is mainly given by the damping of SAWs at very high frequencies along their propagation path. It is also important to use lasers offering high repetition rates with sufficient laser power to reduce the measurement time.

Having obtained the dispersion curve the elastic constants are obtained by solving the so-called inverse problem. Unfortunately the expected signal-to-noise ratio should decrease, first, due to the increasing frequency itself, and second, due to the reduced detection area. As it makes no sense to lose resolution when trying to evaluate small changes in elastic constants, a method was developed to circumvent this type of problem to a certain extent. Makarov *et al.* found that accuracy in determining elastic constants can be increased when combining the dispersion of different acoustic modes.¹⁵

VI. CONCLUSIONS

The STM based acoustic wave techniques SATM and SATS were discussed in experiments and concepts. Due to

the high frequency of acoustic waves in the range from some 10 MHz to some GHz a mixing technique that demodulated the fast signals was employed. Using an IDT for defined Rayleigh wave excitation the phase and the amplitude of surface oscillations were measured. On the scale of crystal grain structures atomic scale steps were seen even though the surface was quite rough, similar to a differential STM. On a Au(111) surface, atomic resolution phase and amplitude images were measured and could be modeled starting from the topography and assuming atomic elliptical oscillation trajectories. For obtaining elastic constants of spatially small structures SATS, a combination of laser pulse excitation of acoustic waves and broadband STM detection was proposed. The initial setup elements realized so far were discussed.

¹K. D. Wise and K. Najafi, *Science* **254**, 1335 (1991).

²R. Kuschneireit, H. Fath, A. A. Kolomenskii, M. Szabadi, and P. Hess, *Appl. Phys. A* **61**, 269 (1995).

³H. Nishino, Y. Tsukahara, Y. Nagata, T. Koda, and K. Yamamaka, *Appl. Phys. Lett.* **62**, 2036 (1993).

⁴P. Hess, *Photoacoustic, Photothermal and Photochemical Processes* (Springer, Berlin, 1989).

⁵J. Kushibiki, T. Ueda, and N. Chubachi, *Proceedings of the 1985 IEEE Ultrasonics Symposium* (IEEE, New York, 1985).

⁶R. Nötzel and K. Ploog, *Int. J. Mod. Phys. B* **7**, 2743 (1993).

⁷C. Thomsen, H. T. Grahm, H. J. Maris, and J. Tauc, *Phys. Rev. B* **34**, 4219 (1986).

⁸W. Rohrbeck and E. Chilla, *Phys. Status Solidi A* **131**, 69 (1992); E. Chilla, H.-J. Fröhlich, J. Riedel, and W. Rohrbeck, *Proceedings of the 8th Tagung Akustik/ 11. Winterschule Mikroakustik* (Phys. Gesellschaft, Berlin, 1989) (in German).

⁹E. Chilla, W. Rohrbeck, H.-J. Fröhlich, R. Koch, and K. H. Rieder, *Appl. Phys. Lett.* **61**, 3107 (1992); *Ann. Phys. (Leipzig)* **3**, 21 (1994).

¹⁰M. R. Freeman and G. Nunes, Jr., *Appl. Phys. Lett.* **63**, 1200 (1993).

¹¹S. Weiss, D. F. Ogletree, D. Botkin, M. Salmeron, and D. S. Chenla, *Appl. Phys. Lett.* **63**, 2567 (1993).

¹²S. J. Stranik and P. S. Weiss, *Rev. Sci. Instrum.* **65**, 918 (1994).

¹³K. Besocke, *Surf. Sci.* **181**, 145 (1987).

¹⁴T. Hesjedal, E. Chilla, and H.-J. Fröhlich (unpublished data).

¹⁵S. Makarov, E. Chilla, and H.-J. Fröhlich, *J. Appl. Phys.* **78**, 5028 (1995).

¹⁶J. Heil, J. Wesner, and W. Grill, *J. Appl. Phys.* **64**, 1939 (1988).

¹⁷D. W. Abraham, C. C. Williams, and H. K. Wickramasinghe, *Appl. Phys. Lett.* **53**, 1503 (1988).

¹⁸T. Hesjedal, E. Chilla, and H.-J. Fröhlich, *Appl. Phys. Lett.* **69**, 354 (1996).

¹⁹K. Takata, T. Hasegawa, S. Hosaka, S. Hosoki, and T. Komoda, *Appl. Phys. Lett.* **55**, 1718 (1989).

²⁰D. Schneider, T. Schwarz, and B. Schultrich, *Thin Solid Films* **219**, 92 (1992).

Investigation on a novel vacuum microelectronic pressure sensor with stepped field emission array

S. H. Xia

Institute of Electronics, Academia Sinica, P.O. Box 2652, Beijing 100080, China

J. Liu

Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

D. F. Cui, J. H. Han, S. F. Chen, and L. Wang

Institute of Electronics, Academia Sinica, P.O. Box 2652, Beijing 100080, China

(Received 9 April 1996; accepted 10 March 1997)

This article investigates a novel vacuum microelectronic pressure sensor in which the cathode emission array is distributed over a stepped cathode base instead of over a conventional flat cathode base. Our computer simulation results indicate that using the stepped field emission array gives the potential of increasing the sensor's sensitivity and expanding its measurement range. Computer simulations are presented, and some experimental results are briefly reported. © 1997 American Vacuum Society. [S0734-211X(97)05304-3]

I. INTRODUCTION

The vacuum microelectronic pressure sensor is a relatively new type of pressure sensor first proposed in 1991.^{1,2} It has the advantages of good temperature stability, radiation hardness, high sensitivity, and tiny size. It can be used to measure not only pressure, but also flow speed, fluid level, acoustic intensity, etc., and it is suitable for use in the sensing system of robot and for work in bad environments.

Based on the structure of the conventional vacuum microelectronic pressure sensor, we propose a new modified structure of the vacuum microelectronic pressure sensor, i.e., a sensor with a stepped cathode emission array. The authors have calculated and compared the emission current of the novel stepped cathode with that of the conventional flat cathode. The calculation results indicate that the new sensor has the advantages of better sensitivity and larger measurement range than the conventional vacuum microelectronic pressure sensor. A stepped cathode emission array was fabricated by the authors. The principle and structure of the new sensor, computer simulations, and experimental results are presented in the following.

II. PRINCIPLE AND STRUCTURE OF THE NEW SENSOR

The structure of the conventional vacuum microelectronic pressure sensor is shown in Fig. 1. It consists of four main parts: (1) a pressure sensing film used as the anode, made of metal or silicon; (2) a cathode, usually employing a field emission array formed by distributing the emission tips over a flat cathode base; (3) an insulator between the cathode and the anode; and (4) a minivacuum chamber between the cathode and the anode. The sensor's working principle is as follows. When a voltage is applied between the cathode and the anode, an electrical field is formed in the vacuum chamber. The field at the emission tip can be very strong at a relatively low anode voltage because of the tiny tip radius and the very small space between the cathode and the anode. Field emis-

sion usually occurs when an external pressure presses against the anode sensing film. The deformation of the anode film, in other words, the distance between the cathode and the anode, varies with external pressure, causing variation of the field strength at the emission tip and resulting in a variation in the cathode emission current. Since field emission is very sensitive to the cathode surface field and the influence of cathode temperature can be ignored in the case of field emission, the vacuum microelectronic pressure sensor inherently has the advantages of high sensitivity, radiation hardness, and temperature stability.

In the conventional vacuum microelectronic pressure sensor, the cathode emission tips are distributed over a flat cathode base, shown in Fig. 1. This structure can be improved by distributing the emission tips over a stepped cathode base, as shown in Fig. 2. In order to explain the improvements, let us first discuss the factors that affect the sensitivity and the measurement range of the vacuum microelectronic pressure sensor.

The sensitivity is mainly affected by two factors, i.e., the deformation characteristics of the pressure sensing film and the emission characteristics of the cathode. As for the sensing film, the thinner the film, the easier the deformation occurs under pressure, and therefore the higher the sensitivity of the sensor is. As for the emission current, the higher the current density, the higher the sensitivity of the sensor is. Increasing the anode voltage, reducing the radius of the cathode emission tip, and decreasing the distance between the cathode and the anode result in an increase of the cathode surface field, and therefore enable an increase of the emission current density and the sensitivity of the sensor.

To obtain optimum sensitivity, the space between the anode sensing film and the cathode emission tips should be as small as possible. The sensing film and the emission tips must not, however, be touched under any circumstances to ensure proper operation of the sensor. Since deformation of the sensing film is affected by external pressure, a small space between the cathode and the anode allows small defor-

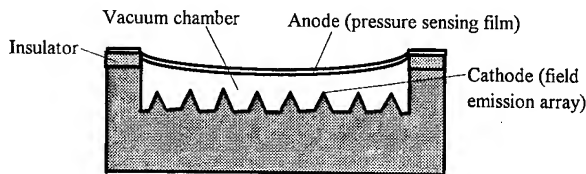


FIG. 1. Principle of the conventional vacuum microelectronic pressure sensor with the emission tips distributed over a flat cathode base.

mation of the sensing film and therefore gives a small measurement range for external pressure. From the above discussion, it is seen that, for the conventional vacuum microelectronic sensor shown in Fig. 1, a contradiction exists in adjusting the cathode–anode space in order to increase the sensitivity and the measurement ranges. The former requires a smaller cathode–anode space while the latter needs a larger cathode–anode distance.

In order to overcome drawbacks of the structure shown in Fig. 1, the authors propose a new structure, shown in Fig. 2, in which the cathode emission tips are distributed over a stepped (concave) cathode base instead of over the conventional flat cathode base. The height and the width of the step are designed according to the deformation characteristics of the anode sensing film to obtain optimum sensitivity as well as a large measurement range.

III. CATHODE EMISSION CURRENT

We now demonstrate the advantages of the stepped cathode over the conventional flat cathode by computer simulations.

The current–pressure relationship, i.e., variation of the emission current with the external pressure, is the main characteristic of the vacuum microelectronic pressure sensor. We have computed and compared the total emission current of the stepped array with that of the conventional flat array at certain anode voltages. The calculation steps are (1) calculate the deformation of the anode film under external pressure to obtain the distance between each emission tip and the anode; (2) calculate the field distribution between the cathode and the anode; and (3) calculate the current of each emission tip as well as the total emission current.

In order to simplify the simulation procedure and reduce the tedious numerical calculation, the authors have assumed that under an external pressure the shape of the anode film is

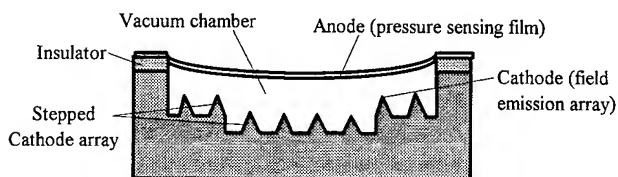


FIG. 2. Principle of the novel vacuum microelectronic pressure sensor with the emission tips distributed over a stepped cathode base.

approximately a parabola, thus simplifying calculation step (1). This simplification does not affect the result of our investigation on the sensor's characteristics.

The potential distribution is computed by solving Poisson's equation $\nabla^2 U = -\rho/\epsilon_0$ using the finite element method (FEM),^{3,4} where U is the potential, ρ is the electron charge density, and ϵ_0 is the permittivity in vacuum. In using the FEM, the corresponding variational functional for axially symmetric systems is

$$F(U) = \iint \{ \epsilon_0 / 2 [(\partial U / \partial z)^2 + (\partial U / \partial r)^2] - \rho U \} 2\pi r \, dz \, dr. \quad (1)$$

The minimization of this variational functional yields the solution of Poisson's equation.

The emission current density J (A/cm²) is calculated by using Child's law, the Schottky formula, and the Fowler–Nordheim formula, respectively, according to the field strength at the cathode surface.⁵

When the cathode surface field is weak and the cathode emission is under space charge limitation, the Child's law is used,

$$J = [4\epsilon_0 (2e/m)^{1/2} V_a^{3/2}] / (9d^2) = 2.33 \times 10^{-6} V_a^{3/2} / d^2. \quad (2)$$

In formula (2), $V_a(V)$ is the potential difference between the cathode and the anode of a diode, and d (cm) is the distance between the cathode and the anode of the diode. A pair of curved equipotential surfaces can be approximately regarded as a diode as long as these two surfaces are very close. In our computer simulation, the cathode surface and its nearby equipotential surface are assumed to form a diode.

When the cathode surface field is in the range of $0 < E < 10^6$ V/cm, the Schottky formula is used,

$$\begin{aligned} J &= AT^2 \cdot \exp(-\phi/kT) \exp[(e^3 E 4\pi\epsilon_0)^{1/2} / kT] \\ &= AT^2 \cdot \exp(-\phi/kT) \exp[0.4403E^{1/2} / T]. \end{aligned} \quad (3)$$

In formula (3), A (A cm⁻² K⁻²) is the material constant of the cathode, T (K) is the cathode temperature, ϕ (eV) is the work function, E (V/cm) is the field strength, and k is the Boltzmann constant.

When the cathode surface field is very strong ($E \geq 10^6$ V/cm), the cathode is under field emission; therefore the Fowler–Nordheim formula is used,

$$\begin{aligned} J &= \{ (e^3 E^2) / [8\pi h \phi \cdot t^2(y)] \} \exp\{ -[8\pi(2m)^{1/2} / (3he)] \\ &\quad \times [(\phi^{3/2} / E)] \theta(y) \}. \end{aligned} \quad (4)$$

In formula (4), $y = (e^3 E / 4\pi\epsilon_0)^{1/2} / \phi$, functions $t(y)$ and $\theta(y)$ have been tabulated. In practice, $t(y)$ and $\theta(y)$ can be approximated as⁶

$$t(y) = 1 + 0.1107y^{1.33}, \quad \theta(y) = 1 - y^{1.69}.$$

In simulating the emission current in the vacuum microelectronic pressure sensor, it is assumed that the total area of the cathode emission array is $300 \times 300 \mu\text{m}^2$, in which 50×50 emission tips are evenly distributed; each tip has a ra-

TABLE I. Comparison of the total emission current of the flat cathode array and the stepped cathode array.

Film deformation	Film voltage (V)	Total emission current (μA)		
		Flat cathode array		Stepped cathode array
		2 μm tip-film space	2.5 μm tip-film space	2.5 μm space at middle 2 μm space at outside
No deformation	350	0.96	0.02	0.75
	400	19.9	0.6	15.4
	450	214.0	9.7	166.9
Minimum t - f space 1.0 μm	350	201	134	317
	400	2 050	1 368	3 164
	450	12 702	8 486	19 209
Minimum t - f space 0.5 μm	350	343	258	416
	400	3 394	2 550	4 084
	450	20 458	15 375	24 433

dus of curvature of 50 nm, material constant of $120 \text{ A/cm}^2 \text{ K}^2$, temperature of 300 K, and work function of 3.5 eV.

For three different sensor structures, the emission currents are calculated at three different anode voltages, 300, 400, and 450 V, respectively. The structures simulated are (1) the flat cathode base, with a 2 μm space existing between the anode film and the cathode emission tips (called the tip-film space) with no external pressure applied (i.e., the anode film has no deformation); (2) the flat cathode base, with a 2.5 μm tip-film space with the anode film having no deformation; (3) the stepped (concave) cathode base, with the middle area lower than the rest of the cathode base; 24 \times 24 tips at the middle area of the cathode base each having a tip-film space of 2.5 μm , and the rest of the emission tips each having a tip-film space of 2 μm . For each structure, the total emission current is calculated at three different anode deformations: (1) without external pressure, i.e., the sensing film has no deformation; (2) the sensing film turns to parabolic under an external pressure, and the distance between the emission tip and the bottom of the parabola is 1.0 μm ; and (3) the sensing film turns to parabolic under an external pressure, and the distance between the emission tip and the bottom of the parabola is 0.5 μm . The calculation results are listed in Table I.

It is seen from Table I that (1) for all the structures simulated (including both the flat cathode array and the stepped cathode array), the emission currents increase with the anode voltages (350, 400, and 450 V). (2) When no external pressure applied, the cathode surface field is relatively weak because of the larger tip-film space, and the emission current is small. (3) When the anode sensing film is deformed by an external pressure, the tip-film space decreases (minimum tip-film space 1.0 and 0.5 μm) and the emission current increases. (4) For the sensor with the flat cathode emission array (tip-film space 2 and 2.5 μm), it is obvious that the smaller the tip-film space, the larger the emission current, and therefore the higher the sensor's sensitivity. (5) For the stepped cathode emission array (2.5 μm tip-film space at the middle and 2 μm tip-film space at the rest of the cathode

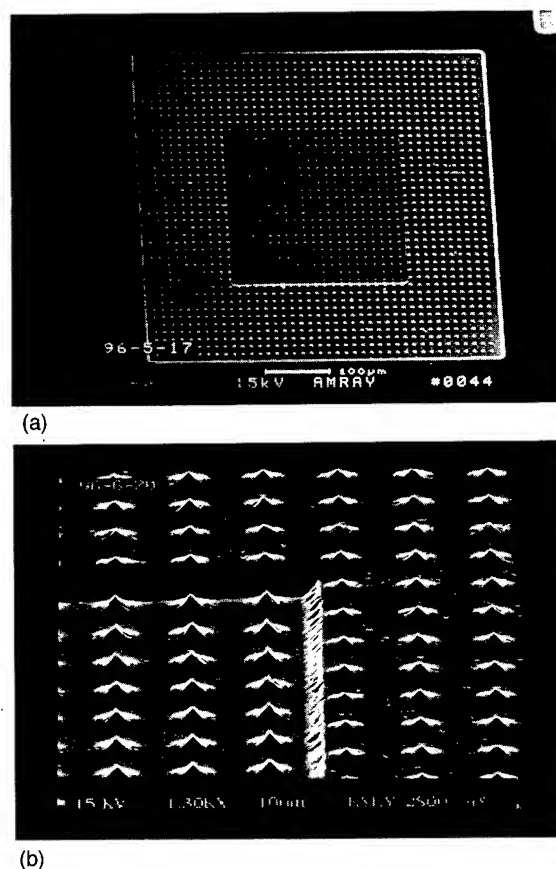


FIG. 3. Scanning electron microscopy photographs of a stepped cathode array: (a) An overall view of the array; (b) part of the array.

base), when no external pressure applied, because of larger tip-film space in the middle and smaller tip-film space in the remaining areas, the total emission current is contributed mainly by the tips distributed over the upper step of the cathode base. Therefore the value of the emission current is close to that of the flat cathode array with the 2 μm tip-film space. (6) When the sensing film is deformed by external pressure (minimum tip-film space of 1.0 and 0.5 μm), the variation of the total emission current of the stepped cathode array is obviously larger than that of the flat cathode array; therefore the conclusion is that the sensor with the stepped cathode array has the highest sensitivity.

The above computer simulation results indicate that the vacuum microelectronic pressure sensor with the stepped cathode array has advantages over the conventional sensor with the flat cathode array with regard to increasing the sensor's sensitivity and expanding its measurement range.

IV. FABRICATION OF THE STEPPED CATHODE ARRAY

Fabrication of the novel stepped cathode array has been realized by the authors. Characteristics of the stepped cathode array are that the surface of the cathode base turns into a concave (stepped) shape, i.e., lower at the middle and higher at the other areas, and the emission tips are distributed over

the upper and lower steps. The height of the step and the size of the lower middle area are designed according to the need.

Figures 3(a) and 3(b) show scanning electron microscopy (SEM) photographs of a stepped cathode array: an overall view of the array and part of the array, respectively. It can be seen that the height of the step is about 5 μm . The emission tips are evenly distributed over the upper and the lower steps, and the radius of curvature of each tip is about 50 nm.

A practical measurement of the emission current from the stepped cathode array has not been performed yet. Experiments are underway to measure and compare the emission currents of the novel stepped cathode array and the conventional flat cathode array.

V. CONCLUSIONS

In this article we have proposed and investigated a new type of vacuum microelectronic pressure sensor in which the cathode emission tips are distributed over a stepped cathode base instead of a conventional flat cathode base. The authors

have computed and compared the emission current of the stepped cathode array to that of the flat cathode array. The computer simulation results indicate that the new sensor has the advantages of higher sensitivity and a larger measurement range over the conventional sensor. Fabrication of the novel stepped cathode array was also realized.

ACKNOWLEDGMENT

The authors acknowledge the financial support of this research from the Chinese Academy of Sciences and the National Natural Science Foundation of China (Project Grant No. 69674030).

¹J. C. Jiang, R. C. White, and P. K. Allen, *Transducers '91*, pp. 238–240.

²H. C. Lee and R. S. Huang, in *Ref. 1*, pp. 241–244.

³S. Xia and A. N. Broers, *Int. J. Numer. Model. Electron. Networks Devices Fields* **8**, 109 (1995).

⁴X. Zhu and E. Munro, *J. Vac. Sci. Technol. B* **7**, 1862 (1989).

⁵S. Xia, Ph.D. dissertation, Cambridge University, UK, April 1994.

⁶P. W. Hawkes and E. Kasper, *Principles of Electron Optics: Applied Geometrical Optics* (Academic, New York, 1989), Vol. 2, p. 927.

Robotic nanomanipulation with a scanning probe microscope in a networked computing environment

C. Baur, B. C. Gazen, B. Koel, T. R. Ramachandran, A. A. G. Requicha,^{a)} and L. Zini
Laboratory for Molecular Robotics, University of Southern California, Los Angeles, California 90089-0781

(Received 12 September 1996; accepted 31 March 1997)

This article describes the initial phase of the development of a high-level programming system for robotic manipulation with a scanning probe microscope (SPM). A SPM server has been developed, which runs in the WINDOWS environment of the PC that controls the microscope. Client programs running on Unix work stations or other computers connected to the Internet can send remote commands to the SPM through the server. The clients can be sophisticated artificial intelligence programs that reason about robotic tasks and sensory data acquired by the SPM. A first set of intermediate-level commands has also been developed for sensing and manipulation. The system is being tested by pushing colloidal gold nanoparticles with dimensions in the order of 15–30 nm on a mica substrate in noncontact atomic force microscope mode. The test programs image the sample, search for nanoparticles in the presence of thermal drift, turn feedback on and off for pushing, and so on. The particles are being moved reliably. © 1997 American Vacuum Society.
[S0734-211X(97)11004-6]

I. INTRODUCTION

The Laboratory for Molecular Robotics at the University of Southern California is an interdisciplinary research laboratory whose ultimate goals are the development of a high-level programming system for robotic manipulation with a scanning probe microscope (SPM), and its use in challenging applications such as building nanomachines or DNA manipulation. Ultimately we expect to be able to process very high level commands such as “assemble object X with object Y,” where X and Y might be, for example, molecules or nanoparticles. Robotic systems at the nanoscale must, like their macroscopic brethren, sense, “think,” and act. The SPM is both a sensor and an actuator. But we must provide it with the intelligence necessary for reliably executing complex commands. This requires sophisticated artificial intelligence programming and software tools that typically run on Unix workstations.

Previous research shows that it is possible to position with SPM atoms, molecules, nanoparticles, and liquid droplets.^{1–4} The reported approaches exhibit one or more of the following characteristics:

- (i) Careful selection of substrates and objects to be moved.
- (ii) Trial-and-error determination of SPM parameters for imaging and manipulation.
- (iii) Operation in ultrahigh vacuum (UHV) and/or low temperature.
- (iv) Time-consuming user-interactive manipulation.
- (v) Custom software developed, typically for an IBM-compatible PC.

In the macroworld, robots typically are programmed by technicians or bachelor-level engineers, and execute reliably

and routinely thousands of tasks per day in industrial environments. In contrast, nanomanipulation today is more of an experimental *tour-de-force*, accomplished in careful experiments run by Ph.D.-level researchers in a few laboratories, than a useful technique that can be used routinely for practical purposes. The work in our laboratory seeks to make nanomanipulation tasks much easier to program and execute reliably, so as to transform nanomanipulation into a mainstream technique, much like its macroscopic counterpart.

Commercially available software for SPMs typically runs on stand-alone PCs. The growth of the Internet and the difficulty of co-locating personnel in interdisciplinary research endeavors (e.g., our own lab is spread across several buildings) both point to the need for distributed software systems. PCs are not the platforms of choice for the development of intelligent systems in a distributed environment, and research institutes tend to have a heterogeneous computing environment, with most of the sophisticated programming being done in Unix work stations. For these reasons, software that makes the SPM behave as a network “peripheral” is very desirable. The SPM becomes thus available to programs running on any computer in the Internet. This also has the advantage of making net resources such as printers and mass storage readily available to SPM users.

Commercial SPM hardware and software are primarily directed at imaging. However, manipulation and imaging have different requirements. One can envisage at least two strategies for developing SPM nanomanipulation facilities. The first consists of designing and building from scratch a custom instrument, together with its low-level and high-level control software for manipulation tasks. This may lead to optimal results but requires substantial resources in terms of time, money, and personnel. An attractive alternative, discussed in this article, starts from commercially available hardware and low-level control software, and builds higher-level layers of software upon them.

^{a)} Author to whom correspondence should be addressed; Electronic mail: requicha@lipari.usc.edu

The remainder of this article is organized as follows. First we describe a SPM server that provides communication facilities for using the SPM transparently in a network. Next we discuss some of the intermediate-level manipulation commands we have developed. Then we present experiments in nanomanipulation that use our software, and conclude with a brief assessment of our current manipulation capabilities and open issues.

II. SPM SERVER

We are using an AutoProbe CP instrument, a commercially available SPM developed by PSI (Park Scientific Instruments). The SPM software is a 16-bit application that runs in IBM-compatible PCs under WINDOWS 3.1 or WINDOWS 95. It includes a modern graphic user interface, through which a user issues commands. These are translated and sent to a digital signal processor (DSP) that controls the instrument. More important, PSI also provides an application programming interface (API) to its software. This API is meant to be used by experienced programmers. We think that working through a commercial API has significant advantages. It lets a programmer issue low-level commands without having to get involved in the intricacies of the DSP and real-time control, which are best left to the manufacturer. It also greatly facilitates software maintenance, since it relies on the manufacturer to ensure that new developments in hardware and low-level software are smoothly incorporated into the system through the API.

In Unix systems a process running in one machine can invoke procedures that reside in a different machine through a remote procedure call (RPC). Unfortunately RPCs are not supported in the WINDOWS 3.1 or WINDOWS 95 environments. To achieve similar functionality one can use a SPM server that listens to requests for SPM operations issued by computers in the network, and takes the appropriate actions. Typical commercial computer communication packages, for example Sun Microsystems' PC-NFS, are aimed at building PC clients, and do not directly support PC servers. Because powerful tools such as *socket* libraries are available, building a server is not a large task for a computer scientist versed in distributed systems. (A socket is a programming abstraction that corresponds to a communication port through which messages can be sent or received.) Furthermore, commercial systems normally do not provide access to source code, which we believe is important for a research group that is exploring new approaches and techniques that may require intimate knowledge of the software. For these reasons, we decided to develop a SPM server, as explained below.

The PC that controls the SPM is connected to our Ethernet Local Area Network through an Ethernet card. The SPM server is a WINDOWS application that uses a standard inter-process communication library called Winsock. The server opens a socket into the network and listens to requests from client processes running on other machines. These clients issue requests by calling a Unix-standard socket library. Both Winsock and the Unix socket library use the standard TCP/IP protocols, but they isolate the programmer from the

low-level details of how communications between computers are achieved. Socket programming is a relatively straightforward exercise. The clients' requests are decoded, mapped into API routines, and appropriate API calls are issued. (Examples of API calls are given in Sec. III.) The results of API routine execution are received by the server, encoded for transmission, and sent back to the requesting process. We encode the messages in external data representation (XDR) format, which is a *de facto* standard introduced by Sun Microsystems.

We wrote a library of C procedures that mirror those provided in the SPM's API, and run on Unix work stations. These procedures exchange information with the PC through the network, by using sockets. A client program running in a Unix work station invokes procedures in this library, and is unaware of the network communications involved. In essence, the messages sent by a client correspond to API calls, and the messages it receives correspond to the results returned by the API routines called. More sophisticated capabilities are being built on the client side, on top of the low-level API, as explained below.

III. MANIPULATION COMMANDS

The graphic user interface supplied by PSI is designed for imaging and is not suitable for manipulation. But the API has powerful routines that can be used to build manipulation commands. An important low-level API procedure serves to apply ramp (i.e., linearly increasing or decreasing) signals to SPM input ports. A ramp is specified by a list of triplets (*InitialValue*, *Slope*, *Duration*). For example, applying two ramps with the same slope and duration to the x and y inputs of the piezoelectric scanner causes a linear motion at 45° to the x axis, with a speed determined by the slope of the ramps, and starting at the specified initial point. As a second example, a step change in a SPM parameter value can be accomplished by applying two successive "ramps" of zero (horizontal) slope, starting at appropriate initial values.

By using these ramps it is easy to construct a fundamental command to move the SPM tip from (x_1, y_1) to (x_2, y_2) at speed s . In turn, this linear motion capability can be used to traverse more complex trajectories. For example, a circular trajectory can be followed by decomposing it into linear segments. More interestingly, imaging scans of arbitrary size and orientation can also be built up with sequences of linear motions.

The API also provides procedures for setting SPM parameter values, turning feedback on and off, and so on. Two of the most useful commands we have implemented thus far are the following:

- (1) Move the tip in a straight line from (x_1, y_1) to (x_2, y_2) at speed s , turn the z feedback off at some prescribed point along the trajectory, and then turn it back on at another prescribed point. This command is called *Push*, since it is used to push nanoparticles.
- (2) Move the tip in straight line from (x_1, y_1) to (x_2, y_2) at speed s and output the z wave form. This is called a

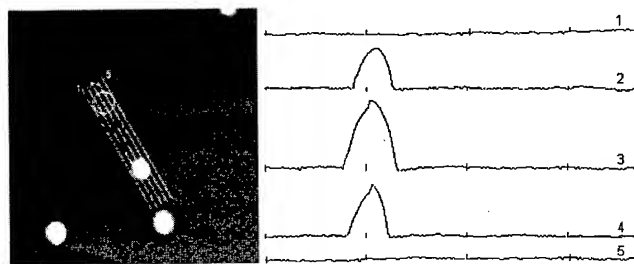


Fig. 1. A particle to be moved and five single line scans in the direction of motion.

SingleLineScan and is used to search for particles, as explained below.

It is important to note that manipulation commands are issued in *instrument* or *robot coordinates*, whereas what we want to accomplish is best expressed in *sample* or *task coordinates*. The robot and task coordinate systems do not coincide, and their relationship is time-dependent. The major cause for this discrepancy is thermal drift. Drift in our SPM is large when the instrument is turned on, but stabilizes at about 1 Å/s after a couple of hours of operation. In addition, if we command the tip to move to a point (x, y) along a straight line, the motion will not be exactly as commanded due to piezo nonlinearities, creep, and hysteresis. Nonlinearity is compensated to some degree by the manufacturer's software, but creep and hysteresis cause significant problems.

IV. PUSHING EXPERIMENTS

We are testing our manipulation software by pushing colloidal gold balls on a mica substrate with the SPM in atomic force microscope (AFM) mode, in air, at room temperature. We have thus far tried Au balls with diameters of 27 and 15 nm, with similar results.

The gold colloidal particles are deposited on a mica substrate through a procedure normally used for the preparation of samples to calibrate AFMs and transmission electron microscopes (TEMs).⁵ The mica surface in water is negatively charged,⁶ as are the gold particles.⁷ A positive coating of the mica surface with Poly-L-Lysine allows the gold particles to be adsorbed onto it.⁸⁻¹⁰

We first attempted to image the Au particles in contact AFM (C-AFM) mode, with no success. The tip seems to disturb the particles significantly. We switched to noncontact AFM (NC-AFM) and obtained consistently good images (or consistently poor images, for bad probes). We tried to push the particles by moving the SPM tip approximately across the particles' centers in C-AFM mode, but failed. We finally managed to move the nanoparticles by turning off the z feedback in NC-AFM mode, without altering the input vibration to the cantilever. (We do not know yet whether the vibration helps or hinders the pushing operation.)

The nanoparticles do not move significantly on their own at the time and spatial scales of the experiment, and can be used to establish a task coordinate system. Successive im-

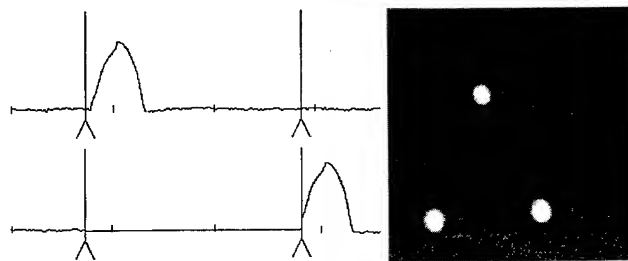


Fig. 2. Single line scans through the particle before and after motion (left) and the resulting particle configuration (right).

ages of the same sample region differ because of the drift between the robot and task coordinate systems. For simplicity, let us assume that initially the robot coordinates (x, y) coincide with the task coordinates (X, Y) . Suppose that we detect a particle at position (x, y) in an image, and want to move it in the X direction by a distance D . The basic strategy is to move the tip to $(x - \Delta x, y)$, where Δx is a small displacement, turn off the feedback, move along x until $x + D$, and turn the feedback on. However, we found experimentally that this often fails, because the robot and task coordinate systems no longer coincide when the command is executed.

This problem can be tackled by either estimating the new position of the particle in the robot coordinate system, or by using updated measurements. We took the latter tack. Through a sequence of quick *SingleLineScan* commands we search for the center of the particle in updated robot coordinates, and use these coordinates to *Push*. Figure 1 shows on the left a particle to be moved to the location marked by the open circle, plus five scan lines used to compute an updated position for the particle. On the right is the z data obtained by the five single line scans. Line 3 is closest to the center of the particle, and selected for the actual motion. Figure 2 shows on the left single scans along line 3 before and after the pushing operation. Feedback is turned on and off at the vertical lines. The right image of Figure 2 illustrates the resulting configuration after pushing.

We have implemented two versions of our manipulation software. One version runs only on the PC through a graphic user interface, and is used primarily to debug code and strategies. A user can pick coordinates with a mouse, draw trajectories graphically, and set points along a trajectory for turning feedback on and off. The other version does not have a graphic user interface, and consists of a suite of routines that can run on remote machines and communicate with the SPM server.

Figure 3 illustrates a manipulation task accomplished with the PC version. It took several hours to position the Au balls to form the "USC" pattern. Clearly, automation is needed if we are to perform more complex tasks routinely.

The basic algorithm for automating positioning operations is the following.

- (1) Scan the sample to obtain an initial image. Set $(x, y) = (X, Y)$, i.e., assume that the robot and task coordinate systems coincide initially.

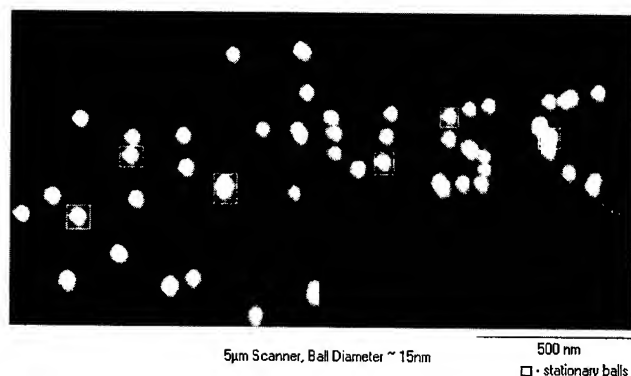


FIG. 3. The random pattern of 15 nm Au balls shown on the left was converted into the "USC" pattern on the right by a sequence of pushing commands.

- (2) Extract from the image the coordinates of the centers of the particles to be moved.
- (3) For each of these particles issue a command to move it in straight line from its current position (X, Y) to the desired position ($X+D_x, Y+D_y$) in task coordinates.
- (4) To implement a move command for a particle, search for the new (x, y) coordinates of its center, by means of *SingleLineScans* as explained earlier, and then issue a *Push* command in robot coordinates.

This simple algorithm assumes that the drift is a pure translation. Otherwise we would have to compute a more complicated relationship between robot and task coordinate systems. It also assumes that all particles will indeed move to their programmed positions. We will return to this latter issue in the Sec. V.

V. DISCUSSION

We have shown how to move Au balls on a mica substrate with a SPM in NC-AFM mode by issuing manipulation commands to a SPM server in a networked computer environment. Our approach uses a commercially available SPM and a manufacturer-supplied API that hides the complexities of real-time control from a higher-level robot programmer. We built intermediate-level commands for sensing and acting on top of this API. We succeed in moving a desired particle an estimated 80%–90% of the time. However, there is much we don't understand yet. Here are some of the issues we are currently grappling with.

- (i) Some particles do not move at all, whereas others move only in a few specific directions, and require suitably chosen intermediate points to reach a desired final position. We do not know why. Particles that are imaged clearly tend to be easy to move.
- (ii) Particle motions sometimes fall short of the commanded distances. With additional sensing and pushing the desired positions can still be reached.
- (iii) Can we characterize in a predictive fashion which particles are movable on which substrates, or do we have to analyze the situations case by case, possibly by empirical methods?
- (iv) Do the strategies described here extend to smaller dimensions, say, to 2 nm particles?
- (v) Do they extend to other particle/substrate pairs? If not, we will not be able to write "generic" pushing software.
- (vi) Can useful devices be built by pushing operations?

ACKNOWLEDGMENTS

The research reported in this article was supported by the Zohrab A. Kaprielian Technology Innovation Fund of the University of Southern California.

¹J. A. Strosio and D. M. Eigler, *Science* **254**, 1319 (1991).

²T. Junno, K. Deppert, L. Montelius, and L. Samuelson, *Appl. Phys. Lett.* **66**, 3627 (1995).

³T. A. Jung, R. R. Schlittler, J. K. Gimzewski, H. Tang, and C. Joachim, *Science* **271**, 181 (1996).

⁴J. Hu, R. W. Carpick, M. Salmeron, and X. Xiao, *J. Vac. Sci. Technol. B* **14**, 1341 (1996).

⁵J. Vesenka, S. Manne, R. Giberson, T. Marsh, and E. Henderson *Biophys. J.* **65**, 992 (1993).

⁶E. A. Hauser, *Silicic Science* (Van Nostrand, Princeton, NJ, 1955).

⁷K. Park, H. Park, and R. M. Albrecht, *Colloidal Gold: Principle Methods and Applications* (Academic, New York, 1989), Vol. 1.

⁸R. Hunter, *Foundations of Colloid Science* (Clarendon, Oxford, UK, 1989).

⁹J. Klein and P. F. Luckham, *Colloids Surf.* **10**, 65 (1984).

¹⁰Fresh monodisperse colloidal particles prepared from standard reduction processes were from Ted Pella, Redding, CA. The Poly-L-Lysine solution (0.1%) was from Sigma Diagnostic. The mica sheets were from Electron Microscopy Sciences, Ft. Washington, PA. The overall procedure consisted of adsorbing 20 μ l of 0.1% Poly-L-Lysine onto freshly cleaved mica for 20–60 s, rinsing with de-ionized water and drying with nitrogen. Immediately after drying, 20 μ l of gold colloidal solution was adsorbed onto the treated mica for 5 min or more, depending on the surface concentration needed. The sample was then rinsed again with de-ionized water. After drying with nitrogen it was finally incubated in a 60 °C oven for at least 1 h.

Using a new kind of organic complex system of electrical bistability for ultrahigh density data storage

H. J. Gao, L. P. Ma, H. X. Zhang,^{a)} H. Y. Chen,^{b)} Z. Q. Xue,^{a)} and S. J. Pang^{c)}
Beijing Laboratory of Vacuum Physics, Chinese Academy of Sciences, Post Office Box 2724,
Beijing 100080, China

(Received 12 September 1996; accepted 31 March 1997)

We present a new kind of organic complex system of electrical bistability, *m*-nitrobenzal malononitrile and diamine benzene (mNBMN-DAB), which can be employed as ultrahigh density data storage devices by scanning tunneling microscopy (STM). The mNBMN-DAB thin film was prepared on highly oriented pyrolytic graphite (HOPG) substrates using physical vacuum deposition method. A critical voltage pulse between the STM tip and the surface of the HOPG substrate can make marks on the thin films. The size of the marks is about 1.3 nm and the corresponding data storage density is larger than 10^{13} bits/cm². A new mechanism of the charge transfer in the system for the data storage is suggested. These results show a great potential of this type of organic composite systems in the application of the ultrahigh density data storage. © 1997 American Vacuum Society. [S0734-211X(97)04704-5]

I. INTRODUCTION

Scanning probe microscopes (SPMs), such as scanning tunneling microscope (STM), atomic force microscope (AFM), scanning capacitance microscopy (SCM), and near-field optics, have been proven powerful to make surface modification on the nanometer scale for high density data storage.¹⁻⁵ In these studies, the entire storage function of recording, erasing, and reproducing was not considered in most studies. Subsequently, it would be extremely important to record and reproduce changes in tunneling current without any surface amorphous change at the surface of the substrate. So far, a few works on the probe-based high density data storage using the phase transition principle,⁶ for example, from amorphous to crystalline, or the phase transition in organometallic thin films, have been reported in the last few years.

Various materials have been used for preparing thin films as storage media, such as Ge₂Sb₂Te₄ and *b*-Na_xV₂O₅ films. Sato *et al.*⁶ suggested that the main mechanism of recording on these films was phase transition from amorphous to crystalline phase. Organic functional complex materials are attractive in nanoelectronics due to their electrical bistability. In this article, we present the high density data storage by STM on a kind of organic complex mNBMN-DAB. The possible mechanism is attributed to the phase transition in the thin films.

II. EXPERIMENT

Two kinds of organic materials, *m*-nitrobenzal malononitrile (mNBMN) and diamine benzene (DAB), were synthesized by our research group. The chemical structures of the mNBMN and DAB can be seen in Fig. 1, respectively. The

two materials were mixed together with nominal composition of 1:1. The mixed *m*-NBMN/DAB powder was deposited on freshly cleaved highly ordered pyrolytic graphite (HOPG) substrates by a vacuum deposition method. At a critical deposition condition, insulating mNBMN-DAB films can be prepared by this method. We employed the mNBMN-DAB films of a few nanometers with the optimum electrical property to obtain desired results. In the mNBMN-DAB system, the mNBMN is an acceptor, and the DAB is a donor.

Experiments were performed with a home-made STM under ambient conditions. The STM tip were 0.25 mm in diameter Pt/Ir (80/20) wire that were snipped with a wire cutter. The STM operated in a constant height mode. Different tips and samples were applied for checking the reliability of the experiments. STM operates in an ambient environment. During imaging the surface of the film, both the constant current mode and the constant height mode were used. A voltage pulse was applied on the thin film between the STM tip and the HOPG substrate. In order to know the electrical property of region on the thin film before and after the voltage pulse, current-voltage (*I*-*V*) relations of the local regions on the thin films were measured by STM.

III. RESULTS AND DISCUSSION

A typical 30 nm×30 nm STM image of the insulating mNBMN-DAB films with recorded marks is shown Fig. 2. A flat region with no marks is selected for recording. The scanning rate is 0.3 Hz/image. The tunneling condition is *V*_b=0.98 V, *I*_t=0.17 nA. The recorded marks are made by applying pulse voltages of 4 V for 10 ms. The recorded marks are stable during the scanning period for more than a few days. The probability of recording is over 90%. The recorded mark is 1.3 nm in diameter, which corresponds to a data storage density of about 10^{13} bits/cm².

Although the STM works in a constant height mode, the feedback circuit is not turned off completely. When the tip is scanning over the recovered marks, the distance between the

^{a)}Present address: Department of Electronics, Peking University, Beijing 100871, China.

^{b)}Present address: Department of Chemistry, Peking University, Beijing 100871, China.

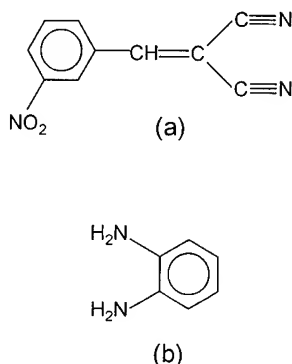


FIG. 1. Chemical structure of (a) mNBMN and (b) DAB.

tip and film becomes larger due to the higher conductance of the marks. It can be seen from Fig. 2 that the recorded marks have dark shadows behind them. This is because the scanning rate and the conductance of the recorded marks are so high that the scanning tip cannot immediately recover its original height level just behind the recorded marks.

It is also of great importance that in our experiment, what is the formation mechanism of the dark marks, surface modification, or phase transition of the thin film from high impedance to low impedance? Sato *et al.*⁶ reported the high density data storage by STM on a composite material, a crystalline and its glassy, and attributed the recording marks to the phase transition of the material from amorphous to crystalline. Mutsumoto *et al.*⁷ used the Cu-TCNQ to carry out high density storage, and believed that the possible mechanism was also the phase transition. However, they did not give further explanations. To understand the formation mechanism of the formed dark marks, $I-V$ curves of the sites were measured by STM before and after the voltage pulses. Before the voltage pulse of the thin film, two typical $I-V$ relations were obtained as shown in curves (a) and (b) of Fig. 3. Curve (c) of Fig. 3 is the $I-V$ relation of the substrate HOPG without sample. In the curve (a), the applied voltage is from 0 to 2.8 V, and the tunneling current is 0 nA when the ap-

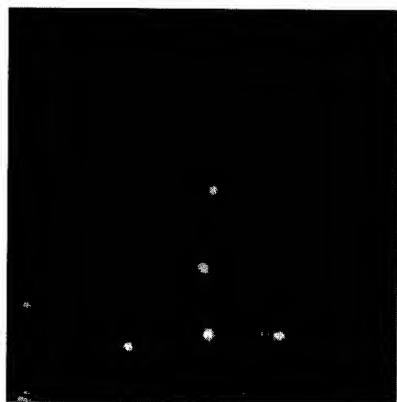


FIG. 2. Data storage on NBMN-DAB organic complex thin film. The size of each marks is about 1.3 nm, and the corresponding density of the data storage can be 10^{13} bits/cm². During the scanning, $V_t = 0.98$ V, $I_t = 0.17$ nA. The pulsed voltage is 5 V and 0.5 ms.

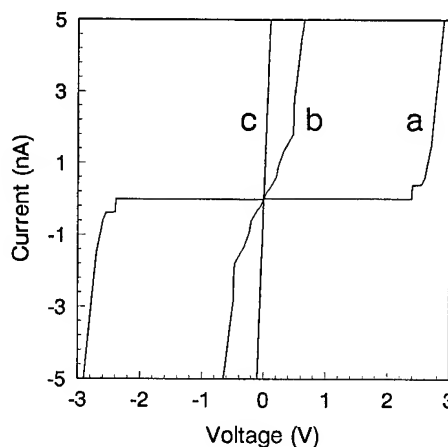
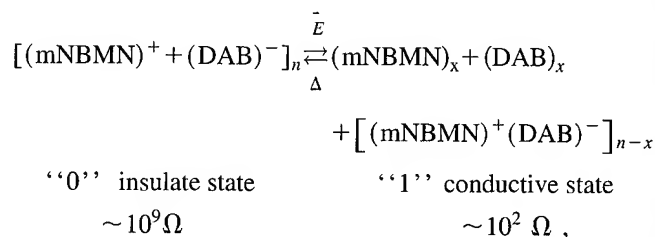


FIG. 3. Typical current-voltage relations of the unrecorded and recorded regimes by STM. Curve (a) shows the $I-V$ relation before the voltage pulse, indicating that the media was initially insulator below the voltage threshold 2.4 V, and became conductive when the voltage excess the 2.4 V. Curve (b) represents the $I-V$ relation of the created regime, indicating that the recorded regime is conductive. Curve (c) is the $I-V$ relation of the HOPG substrate, showing the straight line and the linear $I-V$ relation.

plied voltage is below 2.4 V, which means that the mNBMN-DAB thin film is initially an insulator below the applied voltage of 2.4 V. When the applied voltage increased and exceeded 2.4 V, we can see the $I-V$ relation from curve (a) of Fig. 3 is very different from that in the initial part of curve (a), which suggest that the film was initially an insulator, and after the voltage threshold of 2.4 V, it changed from the insulator to a conductor. After the voltage pulse, the $I-V$ relation of the pulsed sites was shown in the curve (b) of the Fig. 3, which indicates that the pulsed sites of the thin film were conductive.

On the other hand, in the four-probe ($I-V$) measurement of the sample, we previously found the electrical bistability of the thin film.^{8,9} The films of electrical bistability have an electric field threshold. Also, the film is initially in a high impedance state, if the applied electric field is below the voltage threshold, it would maintain the high impedance, and when the applied electric field exceed the threshold, the film would become a conductive state. The film transition can be described as follows:



where Δ represents heating which can recover the recorded regimes from 1 state to 0 state.

Compared the $I-V$ curves in the STM measurement with those in the four-probe $I-V$ measurement, one can find the common point, i.e., both the electric field between the STM tip and the substrate and the two electrodes in the four-probe $I-V$ measurement can cause the transition of the electrical

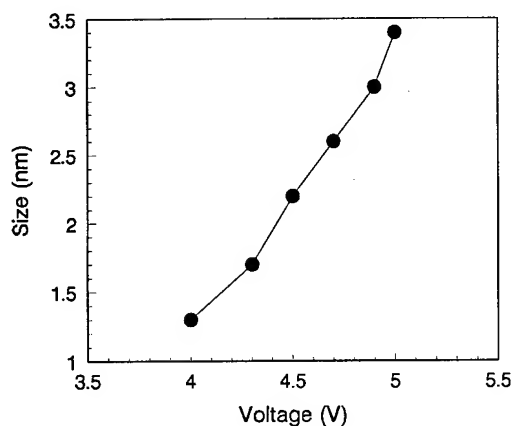


FIG. 4. Relation of the size of recording marks with the amplitude of the applied voltage pulse, indicating that the recording marks could be created on the organic complex by applying 4–5 V voltage pulse for a few microseconds.

property of the mNBMN-DAB thin film. In both cases, the voltage thresholds are almost in the same order, which suggests that the electrical property of the pulsed sites of the thin film changed greatly from insulator to conductor. To further evidence the formation mechanism of the dark sites, we employed the different samples with different thickness and different components of mNBMN and DAB in the thin film, and we found that though the voltage pulse from the STM tip was the same as that described above, no dark mark could be made on the thin films of no electrical bistability. Only the thin films with the electrical bistability and at a suitable voltage pulse from the STM tip can form the dark mark, which further demonstrate that the possible mechanism of the mark formation is caused by the phase transition from a high impedance to a low impedance of the organic composite thin films.

In this recording experiments, different tips and insulating mNBMN-DAB films were used, and it was found that the recording mark could be made on the films by applying voltage pulse. The mark size varied with the amplitude and the width of the voltage pulse. The size of the mark varied with the amplitude and the width of the voltage pulse. At a fixed voltage pulse of 10 ms, the relation between the size of the mark and the amplitude of the voltage pulse is shown in Fig. 4. It was found that recording mark could be made on the organic complexes by applying 4–5 V voltage pulse for several microseconds. Both positive and negative voltage pulses can make the same kind of recording, which means that material evaporation induced by electric field is not the main reason for the mark formation. We tried to apply a voltage pulse to some conducting mNBMN-DAB films, but no recording mark could be made. Therefore the recording is an intrinsic phenomenon of the insulating mNBMN-DAB films. In the experiment, heating can make the recorded films back to insulator again.

In our STM investigation of the ultrahigh data storage, it should be noted that, the voltage bias should be set below a certain value at the initial scanning of the sample. Otherwise, the film would transfer from a high impedance to a low impedance, and it would become impossible for the film to conduct the data storage. So in the STM studies, we initially apply a low voltage to prevent the film to become low impedance (conductor), in other words, to hold the high impedance of the film in the 0 state. From the above description, it is reasonable to induce that it is the conductance transition of the mNBMN-DAB thin films that makes the possibility of ultrahigh density data storage on the charge-transfer organic composite thin film by STM. However, much more research work should be paid in the future on this kind of materials for the UHDDS by STM both experimentally and theoretically.

IV. CONCLUSION

We have found a new kind of physical system of electrical bistability, nitrobenzal malononitrile (NBNM) and diamine benzene (DAB) composite thin film, to carry out the ultrahigh density data storage. The storage density is larger than 10^{13} bits/cm². The possible mechanism of the mark formation in the STM image is attributed to the phase transition of the thin film conductance before and after the voltage pulse on the organic thin films. Our results indicate that the employment of STM to the functional organic composite thin films of electrical bistability may show a great potential in the application of ultrahigh density data storage. However, in the future, much more research work should be paid on controlling over the structural properties, synthesizing more stable and reproducible organic complexes, the write/erasure mechanism of the electrical bistability in the organic complex system as well as the effective controlling over the scanning probe microscopy (SPM).

ACKNOWLEDGMENTS

The authors would like to thank R. Colton for helpful discussion and suggestions, Y. Song for help with synthesizing the organic raw material, and Z. Bian for doing some STM experiments. This research is supported in part by the National Science Foundation and the Doctoral Program Foundation of Higher Education Commission of China.

¹R. S. Becker, J. A. Golovchenko, and B. S. Swartzentruber, *Nature* (London) **325**, 419 (1987).

²C. F. Quate, US Patent No. 4, 575 (1986).

³R. C. Barrett and C. F. Quate, *Ultramicroscopy* **42**, 262 (1992).

⁴H. J. Mamin and D. Rugar, *Appl. Phys. Lett.* **61**, 1003 (1992).

⁵H. J. Mamin, B. D. Terris, S. Hoen, and D. Rugar, *IBM J. Res. Dev.* **39**, 681 (1995).

⁶A. Sato and Y. Tsukamoto, *Adv. Mater.* **6**, 79 (1994).

⁷M. Matsumoto, Y. Nishio, H. Tachibana, T. Nakamura, Y. Kawabata, H. Samura, and T. Nagamura, *Chem. Lett.* 1021 (1991).

⁸H. J. Gao, Z. Q. Xue, Q. D. Wu, and S. J. Pang, *J. Vac. Sci. Technol. B* **13**, 1242 (1995).

⁹H. J. Gao, Z. Q. Xue, K. Z. Wang, Q. D. Wu, and S. J. Pang, *Appl. Phys. Lett.* **68**, 2192 (1996).

Atomic force microscope-based data storage using replicated media

B. D. Terris

IBM Research Division, Almaden Research Center, San Jose, California 95120-6099

S. A. Rishton

IBM Research Division, T. J. Watson Research Center, Yorktown Heights, New York 10598

H. J. Mamin, M. E. Best, J. A. Logan, and D. Rugar

IBM Research Division, Almaden Research Center, San Jose, California 95120-6099

(Received 12 September 1996; accepted 20 March 1997)

We have developed a technique for both mastering and replicating data patterns for potential use in an atomic force microscope (AFM)-based data storage device. The process consists of using electron beam lithography to write data features as small as 50 nm and a photopolymerization process to faithfully replicate the written marks. The replicas can be read using a contact-mode AFM tip on a rotating disk, and no change in the signal is seen after 12 days of continuous reading.

© 1997 American Vacuum Society. [S0734-211X(97)10504-2]

I. INTRODUCTION

The application of the atomic force microscope (AFM) and various other scanning probe devices to high density data storage has received much attention lately, and numerous schemes for both writing and reading have been investigated.¹ While data densities on the order of 10–100 Gbits/cm² have been demonstrated, many challenges remain in developing a scanning probe memory device. For contact-mode AFM readback, where the data rate is limited by the lever resonant frequency, progress is being made in developing higher frequency levers and a data rate equal to that of the original CD-ROM (1.2 Mbits/s) has been demonstrated.² While various writing schemes have been demonstrated, most of them are still quite slow and are often not erasable. An alternate approach to the tip-based data writing is to use other high resolution, often very expensive, techniques such as electron beam writing to produce a master disk. The master would then be replicated and read using the compact, inexpensive AFM-based disk drive. This would be a read-only device, similar to the CD-ROM, but would have 100 times higher data density. For such an approach to be feasible, an inexpensive method of replicating the high data density disks, containing nanometer scale features, must be developed.

Recently, methods for fabricating nanometer-scale structures over macroscopic areas have received considerable attention. In particular, various printing methods have been developed where a master is used to pattern a substrate. For example, Whitesides and collaborators³ have patterned self-assembled monolayers on a variety of surfaces and St. John and Craighead⁴ have used octadecyltrichlorosilane as an ink to pattern silicon and metal oxide substrates. Chou and co-workers⁵ have molded 25 nm features in polymethylmethacrylate (PMMA) and then transferred them into silicon using reactive ion etching (RIE). All of these methods are able to faithfully replicate nanometer-scale features in a simple and inexpensive way.

II. SAMPLE FABRICATION

One method of producing nanometer-scale replicas suitable for use in an AFM data storage device is to use a photopolymerization (2P) process. As shown in Fig. 1, an electron beam is used to write a master consisting of bits on a track pitch of 100 nm (the tracks run vertically). The bits are written in a CD-like data pattern and are in increments of 100 nm in length. The smallest mark size is 50 nm×100 nm and this corresponds to a data density of 10 Gbits/cm². This is, of course, is not at the resolution limit of direct electron beam writing. The substrate is a silicon wafer on which approximately 500 nm of oxide is grown. A 100-nm-thick layer of PMMA is spun onto the silicon oxide and baked at 175 °C. The pattern to be replicated is exposed using a vector-scan electron beam machine (IBM VS6) at 50 keV and then developed in a 2:1 mixture of isopropyl alcohol and methyl isobutyl ketone. The pattern is transferred to the silicon oxide by RIE in CHF₃ for approximately 3 min, producing a 60-nm-deep etch pattern. The CHF₃ pressure is 30 mTorr, the rf power density is approximately 0.4 W/cm², and the dc bias is 300 V. Any remaining PMMA is then removed in an oxygen plasma RIE. The SiO₂ is then either used directly as a master, or can be coated with a thin release layer (an amorphous TbFeCo alloy) prior to replication.

This pattern in the silicon oxide master is transferred to a replica using the 2P process.^{6–8} This process has previously been shown to faithfully reproduce patterns with a pitch as small as 600 nm.⁹ As shown here, this can be extended down to written feature sizes of 50 nm, the smallest features we have written. A thin layer of an acrylate based photopolymer is formed between the SiO₂ master and a glass substrate by capillary forces. An adhesion promoter is spun onto the glass substrate prior to application of the photopolymer. The 2P is cured by exposure to ultraviolet (UV) light (365 nm) for 2 min, after which the master is separated from the 2P coated glass substrate. To further crosslink and harden the 2P, the replica is baked in a nitrogen atmosphere at 180 °C for 1.5 h.

The master and replica are compared using both scanning

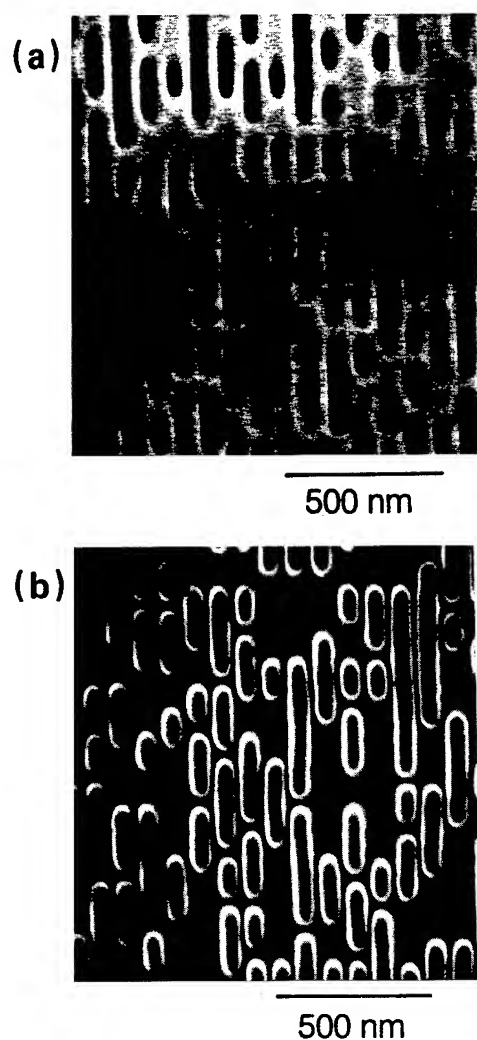


FIG. 1. SEM images of the master (a) and replica (b) of electron beam written features on a track pitch of 100 nm. The minimum mark length is 100 nm and the data density would be 10 Gbits/cm², or 100 times greater than current CD-ROM disks.

electron microscopy (SEM) and AFM. As shown in the SEM images in Fig. 1, the 2P process appears to faithfully reproduce the features of the SiO₂ master, except for a possible slight broadening of the feature width. In the AFM images of Fig. 2, the replicated marks definitely appear broader than those in the master and also higher than the apparent depth of the master marks. Much of this apparent width increase is, however, due to an AFM imaging artifact. The measured width of a bump will be the sum of the actual width and the tip width, whereas the measured width of a pit will be the actual width minus the tip width. From the line scans in Fig. 2, it is also apparent that the tip is unable to image to the bottom of the narrow marks in the master, and the triangular shape of the line scan is indicative of imaging the tip and not the actual pit. The line scans in the replica appear much more representative of the sample and the bump height is in excellent agreement with the known etch depth of 60 nm. Except for the slight broadening of the feature sizes, the 2P process is capable of reproducing marks down to 50 nm in size and

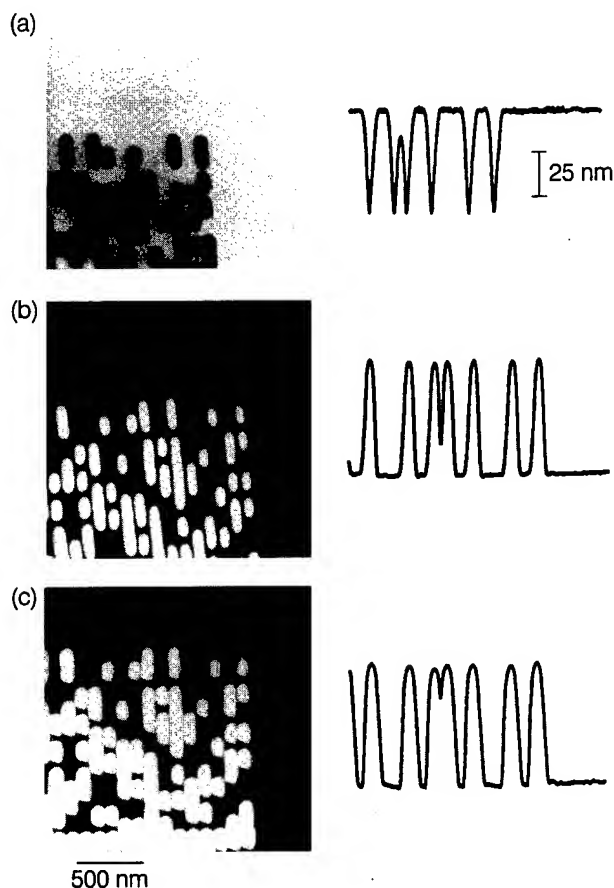


FIG. 2. AFM images of the master (a), a replica cured at 180 °C (b), and a replica held at room temperature (c). The right side of each image is an AFM line scan taken from the top row of marks.

periodic patterns down to 100 nm in pitch. The slight broadening could easily be taken into account when producing a master for data storage applications.

This slight broadening of the feature size can also be controlled by processing conditions. Shown in Fig. 2 are AFM images of 2P replicas, one of which is cured at 180 °C for 1.5 h and one of which is held at room temperature. The room temperature-cured sample shows noticeable broadening and the bump heights are a few nm higher than those in the 180 °C-cured sample. The additional crosslinking in the 180 °C-cured sample clearly results in a more faithful replication of these small features. It is possible the process could be further refined to produce even higher fidelity reproductions.

The replica can also be used as the stamper and a second generation replica produced. In this case, it is necessary to coat the first generation replica with a release layer of amorphous TbFeCo prior to its use as a stamper. As shown in Fig. 3, the second generation replica closely resembles the master. Now, since both samples contain pits, any imaging artifacts from the tip are the same in the two samples, and the pit widths are in good agreement. Thus, if it is desired that the media contain data pits and not bumps, as might be the case for AFM-based readout, this can be accomplished by using a first generation replica as the stamper.

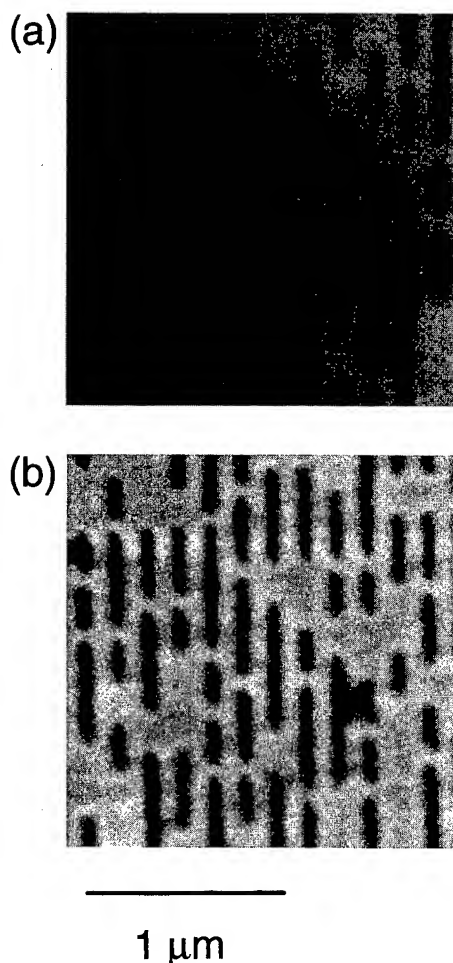


FIG. 3. AFM images of a master and a second generation 2P replica.

In addition to requiring high quality reproduction of written data features, the replication process must not introduce unwanted features or surface roughness. Such roughness will be a source of noise in any proposed data storage readback scheme. The surface roughness measured by AFM is on the order of 0.2–0.5 nm rms for both the SiO₂ master and the 2P replica. This level of roughness is below that of the polycarbonate substrates previously used for AFM data storage experiments in which an acceptable signal-to-noise level was achieved.¹

III. WEAR TESTING

One issue for tip-based data storage is the robustness of both the tip and media when the data are read using a contact-mode AFM. To investigate this, a 2P replica sample is mounted on a rotating air-bearing spindle, and the data read continuously for 12 days. A piezoresistive lever with a Si tip¹⁰ is used and the data are detected as the change in piezoresistance as the lever is deflected by the data bumps. Since the tip and media wear rates will depend on the tip load,¹ the load (approximately 10^{-7} N) is held constant using servo control. This requires a servo bandwidth of at least several times the rotation frequency. With the tip mounted at

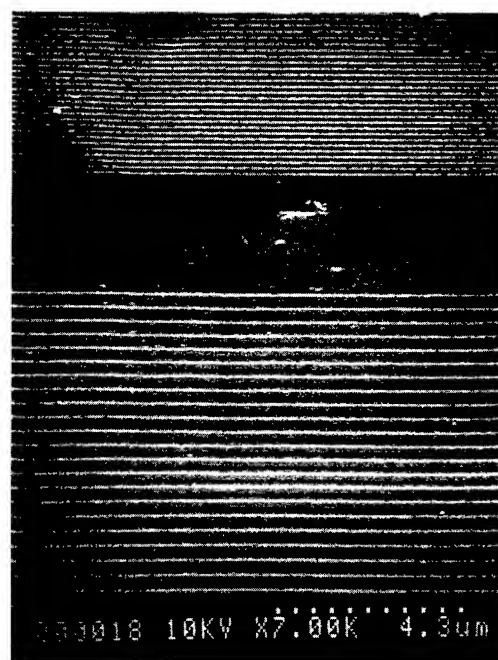


FIG. 4. SEM image of two of the gratings on the 2P replica sample.

a 0.5 mm radius from the center of rotation, the maximum rotation frequency is around 2 Hz, as a higher frequency results in the data rate exceeding the lever's natural resonant frequency. This servo control is achieved by mounting the lever on a CD actuator, after removing the lens. The mass of the lever and holder is approximately equal to that of the removed lens and the tip is roughly centered in the lens

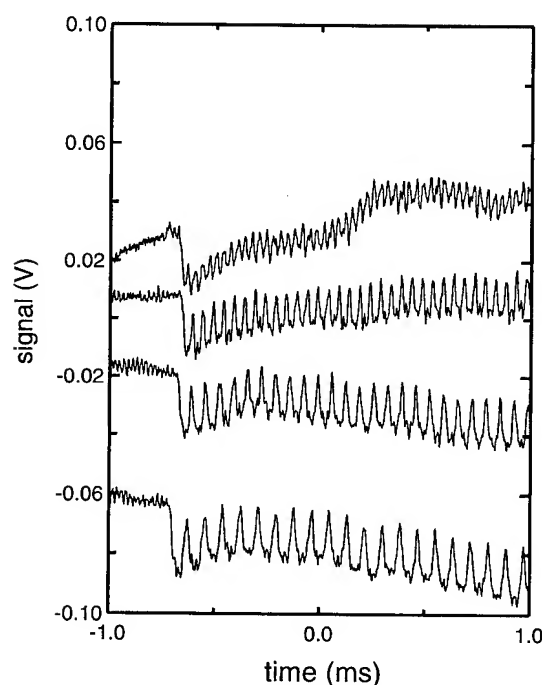


FIG. 5. Piezoresistance signal from the four gratings taken while spinning at 6.25 mm/s. From top to bottom, the grating periods are 200, 300, 400, and 600 nm. The curves are offset vertically for clarity.

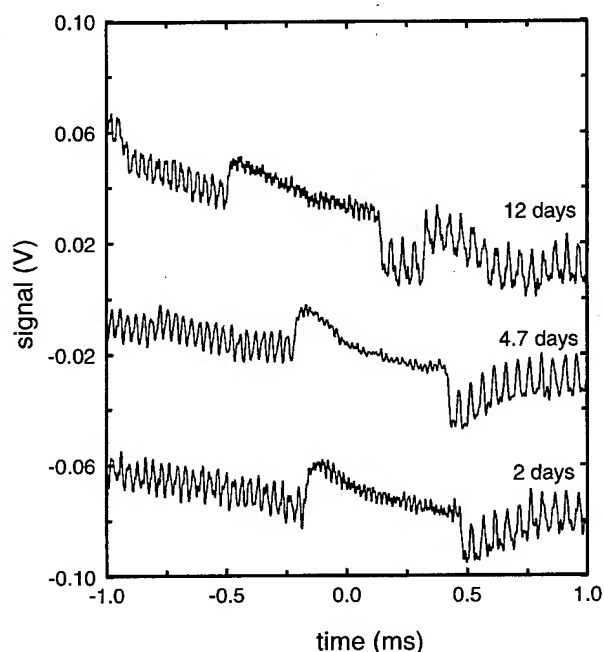


FIG. 6. Piezoresistance signal at three points during the life test. The region on the left side is from the 200 nm period grating and that on the right side is from the 300 nm period grating.

holder. With this, the open loop response falls to a unity gain at around 300 Hz, more than sufficient for the 2 Hz rotation frequency. The servo response of the actuator may not be fully optimized, as the same actuator does not reach unity gain until above 1 kHz when used in an optical servo system.

The 2P replica consists of mostly smooth, featureless regions, with short bursts of periodic linear gratings. There are four gratings with a short smooth region between each one. A SEM image of part of two of the regions is shown in Fig. 4. The periods of the grating are 600, 400, 300, and 200 nm and the depth of the trenches, measured using a silicon nitride tip on a commercial AFM, are 13, 10, 8, and 6 nm, respectively. Shown in Fig. 5 is the piezoresistance signal from each region. The values of the amplitudes and periods are in good agreement with the expected values.

The sample was read continuously for 12 days. As shown in Fig. 6, the signal is essentially unchanged over that time period. This time period represents over 2×10^6 reads, or a tip travel distance of 9.75 km. (The rotation frequency was 10 Hz for the first 43 h.) In this time we saw no degradation of the signal, indicating the tip was still capable of reading a

200 nm period grating. After 12 days, the sample was imaged by optical microscopy (both bright and dark fields) and SEM and no indication of wear was seen. However, since there is no tracking control, the tip quite likely drifts over an area equivalent to many tracks. It is also possible that there are wear tracks, but they are narrow enough that we cannot find them in the relatively wide gratings.

IV. CONCLUSION

In summary, by using mastering and replication techniques, such as those presented here, it is possible to produce high data density replicas that could be used in an AFM-based read-only memory device. In addition, initial life testing suggests that it may be possible to read such disks for long periods of time using contact-mode AFM. At the densities demonstrated, a full CD-ROM of today could be contained on a 10-mm-diam "AFM-ROM" disk. Producing these complex patterns of 100 nm size marks, in circular patterns, over a 10 mm size disk, with accurate mark placement and field stitching, will push the state of the art of electron beam writing. While such a master may be expensive to produce, by using an inexpensive replication process such as 2P combined with batch fabricated AFM tips and levers, a low cost, extremely high capacity form of data storage may be possible.

ACKNOWLEDGMENTS

The authors thanks W. McChesney for the release layer coatings and R. Ried and L.-S. Fan for helpful discussions. The work of two of the authors (B.D.T. and H.J.M.) was partially supported by ARPA Contract No. DABT63-95-C-0019.

¹H. J. Mamin, B. D. Terris, L. S. Fan, S. Hoen, R. C. Barrett, and D. Rugar, *IBM J. Res. Dev.* **39**, 681 (1995).

²H. J. Mamin, L. S. Fan, S. Hoen, and D. Rugar, *Sensors Actuat. A* **48**, 215 (1995).

³S. Palacin, P. C. Hidber, J.-P. Bourgoin, C. Miramond, C. Fermon, and G. M. Whitesides, *Chem. Mater.* **8**, 1316 (1996), and references therein.

⁴P. M. St. John and H. C. Craighead, *Appl. Phys. Lett.* **68**, 1022 (1996).

⁵S. Y. Chou, P. R. Krauss, and P. J. Renstrom, *Science* **272**, 85 (1996).

⁶K. Compagnon and P. Kramer, *Philips Tech. Rev.* **33**, 178 (1983).

⁷H. C. Haverkorn, P. E. J. Legierse, and G. E. Thomas, *Philips Tech. Rev.* **40**, 287 (1982).

⁸J. E. Klooserboer, G. J. M. Lippits, and H. C. Meinders, *Philips Tech. Rev.* **40**, 298 (1982).

⁹M. E. Best and R. B. Prime, *Proc. SPIE* **1774**, 169 (1993).

¹⁰Park Scientific Inc., Sunnyvale, CA.

AUTHOR INDEX

To papers presented at the Fourth International Conference on
Nanometer-Scale Science and Technology

- Abe, Masayuki-(4) 1512
Ahmed, H.-(4) 1402
Arnold, W.-(4) 1506
Ascoli, Cesare-(4) 1398
- Bai, C. L.-(4) 1343, 1350
Bai, Chunli-(4) 1378
Bai, Yubai-(4) 1442
Bandow, Shunji-(4) 1338
Baschieri, Paolo-(4) 1398
Baur, C.-(4) 1577
Behm, R. J.-(4) 1373
Behrendt, F.-(4) 1373
Beltram, Fabio-(4) 1398
Best, M. E.-(4) 1584
Bode, M.-(4) 1330
Bonnell, Dawn A.-(4) 1483
Borusik, O.-(4) 1445
Bösch, M.-(4) 1564
Brillert, Ch.-(4) 1564
- Cao, Yaan-(4) 1442
Cao, Yunwei-(4) 1425
Cecconi, Ciro-(4) 1398
Cháb, V.-(4) 1445
Chai, Xiangdong-(4) 1425
Chen, D. M.-(4) 1295
Chen, H.-(4) 1456
Chen, H. Y.-(4) 1581
Chen, S. F.-(4) 1573
Chen, Sheng-fu-(4) 1385
Chen, W.-(4) 1402
Chen, X. L.-(4) 1502
Chen, Yongmei-(4) 1442
Cheng, H. K.-(4) 1498
Chilla, E.-(4) 1569
Chu, Jiaru-(4) 1551
Chvoj, Z.-(4) 1445
Cui, D. F.-(4) 1573
- De Feyter, S.-(4) 1419
Deng, Huihua-(4) 1460
De Schryver, F. C.-(4) 1419
Ding, Weiping-(4) 1347
Dobisz, E.-(4) 1369
Du, Y. W.-(4) 1456
Du, Youwei-(4) 1347, 1465
- Eguchi, K.-(4) 1429
Esashi, M.-(4) 1414
Esashi, Masayoshi-(4) 1531
Escamilla, Hector-(4) 1516
- Fang, Jinghuai-(4) 1468
Farnault, E.-(4) 1539, 1556
Flörsheimer, M.-(4) 1564
Frediani, Carlo-(4) 1398
Fröhlich, H. J.-(4) 1569
Fuchs, H.-(4) 1564
Fujii, Toru-(4) 1494
Fujisawa, Satoru-(4) 1479
Fujita, H.-(4) 1556
- Gao, H. J.-(4) 1581
Gao, Juning-(4) 1313
Gao, P. J.-(4) 1502
Gastev, S. V.-(4) 1471
- Gazen, B. C.-(4) 1577
Göbel, H.-(4) 1359
Gorwadkar, Sucheta M.-(4) 1406
Grim, P. C. M.-(4) 1419
Gruber, A.-(4) 1382
Gspann, J.-(4) 1382
Gu, Gang-(4) 1347
Gu, Jianhua-(4) 1468
Gu, Minming-(4) 1325
Guo, Zaibing-(4) 1347
- Haak, M.-(4) 1521
Hamanaka, H.-(4) 1414
Han, J. H.-(4) 1573
Hao, H. W.-(4) 1498
Haraichi, Satoshi-(4) 1406
Hartmann, U.-(4) 1364
Hasegawa, Y.-(4) 1300, 1307
Hasunuma, R.-(4) 1437
He, Huixin-(4) 1425
He, Yuliang-(4) 1313
Herrmann, A.-(4) 1521
Hesjedal, T.-(4) 1569
Hiroshima, Hiroshi-(4) 1406
Hirsekorn, S.-(4) 1506
Hodel, U.-(4) 1364
Hoummady, M.-(4) 1539, 1556
Hu, Jun-(4) 1385
Hua, Z. Y.-(4) 1353
Hui, Zheng-(4) 1452
- Ishii, Kenichi-(4) 1406
Itoh, Toshihiro-(4) 1551, 1559
- Jacobs, L.-(4) 1359
Jiang, Yueshan-(4) 1313
Jiang, Yueshun-(4) 1425
Jin, Haiyan-(4) 1442
Jing, Xiao Gong-(4) 1452
- Kaupp, G.-(4) 1521
Kawakatsu, H.-(4) 1539, 1556
Kiely, James D.-(4) 1483
Kilpelä, O.-(4) 1471
Kleiber, M.-(4) 1330
Klyachko, D.-(4) 1295
Kobayashi, Katsuyoshi-(4) 1479
Koel, B.-(4) 1577
Komeda, T.-(4) 1437
Komiya, Masaharu-(4) 1325
Koops, H. W. P.-(4) 1369, 1535
Köster, P.-(4) 1527
Kuroda, R.-(4) 1429
- Lazzarino, Marco-(4) 1398
Lee, Chengkuo-(4) 1551, 1559
Li, L.-(4) 1300, 1307
Li, Lin Song-(4) 1452
Li, Min-qian-(4) 1385
Li, Tie Jin-(4) 1452
Li, Tiejin-(4) 1425, 1442
Li, W. G.-(4) 1350
Li, Xiaodong-(4) 1378
Liang, Bingjie-(4) 1432
Liu, J.-(4) 1502, 1573
Liu, Ning-(4) 1313
Liu, Zhongfan-(4) 1425
Logan, J. A.-(4) 1584
- Löhndorf, M.-(4) 1330
Lu, J.-(4) 1456
Lu, Ran-(4) 1425
Lu, Xiaomei-(4) 1468
Lu, Zuhong-(4) 1460, 1468
- Ma, L. P.-(4) 1581
Machorro, Roberto-(4) 1516
Maeda, Ryutaro-(4) 1559
Mamin, H. J.-(4) 1584
Manassen, Y.-(4) 1310, 1317
Mao, Haifang-(4) 1460
Mao, Haitao-(4) 1465
Masuzawa, T.-(4) 1556
Matsuda, H.-(4) 1429
Memmert, U.-(4) 1364
Miyashita, Masayuki-(4) 1547
Mizutani, W.-(4) 1388
Moessner, G.-(4) 1419
Morita, Seizo-(4) 1479, 1512, 1543
Motomatsu, M.-(4) 1388
Mukaida, H.-(4) 1437
Müllen, K.-(4) 1419
- Nakagiri, Nobuyuki-(4) 1394, 1547
Nakagiri, T.-(4) 1429
Nakashima, Hiroshi-(4) 1411
Nakayama, Kan-(4) 1494
Ni, G.-(4) 1456
Nie, H. Y.-(4) 1388
Nieto, Jesús-(4) 1516
Novikov, S. V.-(4) 1471
- Ohashi, Takahiro-(4) 1559
Ohshima, Kazunari-(4) 1338
Ohta, Masahiro-(4) 1512, 1543
Okada, K.-(4) 1449
Olami, Z.-(4) 1317
Ono, T.-(4) 1414
Ono, Takahito-(4) 1531
Ouyang, M.-(4) 1304
Ouyang, Q.-(4) 1449
Ouyang, Zhen-qian-(4) 1385
- Pang, S. J.-(4) 1502, 1581
Pang, Shijin-(4) 1313
Pascal, R.-(4) 1330
Pingue, Pasqualantonio-(4) 1398
- Qiang, D.-(4) 1304
Qiao, H.-(4) 1498
Qiu, Xiaohui-(4) 1378
- Rabe, U.-(4) 1506
Ramachandran, T. R.-(4) 1577
Ramesh Rao, N.-(4) 1317
Rauscher, H.-(4) 1373
Requicha, A. A. G.-(4) 1577
Rishton, S. A.-(4) 1584
Rücker, M.-(4) 1419
Rugar, D.-(4) 1584
- Saitoh, Hiroaki-(4) 1531
Sakai, Masahiro-(4) 1338
Sako, Sanshiro-(4) 1338
Sakurai, T.-(4) 1300, 1307
Sang, H.-(4) 1456
- Sasaki, Naruo-(4) 1479
Scherer, V.-(4) 1506
Schössler, C.-(4) 1535
Schwarz, U. D.-(4) 1527
Shachal, D.-(4) 1310
Shang, G. Y.-(4) 1343
Shang, Guangyi-(4) 1378
Shi, D. X.-(4) 1502
Shi, Dongxia-(4) 1313
Shido, S.-(4) 1429
Shinohara, H.-(4) 1300
Sigematsu, Tatsuhiko-(4) 1543
Sinkkonen, J.-(4) 1471
Siqueiros, Jesús-(4) 1516
Siqueiros, Jesús M.-(4) 1474
Suga, Tadatomu-(4) 1551, 1559
Sugawara, Yasuhiro-(4) 1479, 1512, 1543
Sugimura, Hiroyuki-(4) 1394
Sugimuraand, Hiroyuki-(4) 1547
Sun, Jie-lin-(4) 1385
Sun, Yue-(4) 1465
Suzuki, Masatoshi-(4) 1494
Suzuki, Mineharu-(4) 1543
Suzuki, Yoshihiko-(4) 1547
- Takimoto, K.-(4) 1429
Ter Ovanessian, E.-(4) 1317
Terris, B. D.-(4) 1584
Tian, F.-(4) 1343, 1350
Tokumoto, H.-(4) 1388, 1437
Tsukada, Masaru-(4) 1479
- Uchihashi, Takayuki-(4) 1512, 1543
Ueyama, Hitoshi-(4) 1512
Uozumi, Kiyohiko-(4) 1411
Urban, J.-(4) 1369, 1535
- Valiyaveetil, S.-(4) 1419
Vanoppen, P.-(4) 1419
von Blanckenhausen, P.-(4) 1359, 1382
- Wada, Toshimi-(4) 1406
Wadas, A.-(4) 1330
Wan, M. X.-(4) 1350
Wang, B.-(4) 1449
Wang, C.-(4) 1343, 1350
Wang, Chen-(4) 1378
Wang, K. Z.-(4) 1304
Wang, L.-(4) 1573
Wang, Li Jun-(4) 1452
Wang, N. X.-(4) 1343
Wang, Shu-(4) 1474
Wang, Yongqiang-(4) 1425
Watanabe, Kazutoshi-(4) 1551
Wei, W. W.-(4) 1498
Wei, Yu-(4) 1432
Weiss, D.-(4) 1330
Wierschem, M.-(4) 1564
Wiesendanger, R.-(4) 1330, 1527
Wu, Jingwen-(4) 1468
Wu, Q. D.-(4) 1304
- Xia, S. H.-(4) 1573
Xiao, Mufei-(4) 1474, 1516
Xu, Hongqi-(4) 1335
Xu, Huijun-(4) 1460
Xu, Jianfeng-(4) 1465

- | | | | |
|------------------------------|-------------------------------|------------------------|--------------------------|
| Xu, W.-(4) 1353 | Yamamoto, Takuma-(4) 1547 | Yuan, Chunwei-(4) 1432 | Zhao, Yong Fang-(4) 1452 |
| Xue, Z. Q.-(4) 1304, 1581 | Yanase, Yoshio-(4) 1543 | | Zhong, Wei-(4) 1347 |
| Xue, Zengquan-(4) 1313 | Yang, Haiqiang-(4) 1313 | | Zini, L.-(4) 1577 |
| | Yang, Jihua-(4) 1442 | Zhang, H. X.-(4) 1581 | Zwörner, O.-(4) 1527 |
| Yahiro, T.-(4) 1539 | Yang, Wensheng-(4) 1425, 1442 | Zhang, J. R.-(4) 1456 | |
| Yamada, Hirofumi-(4) 1494 | Yasuda, S.-(4) 1429 | Zhang, Ning-(4) 1347 | |
| Yamaguchi, Masataka-(4) 1494 | Yi, W. K.-(4) 1498 | Zhang, Y. Z.-(4) 1502 | |

CUMULATIVE AUTHOR INDEX

All authors published to date in Volume 15 are listed alphabetically with the issue (in parentheses) and page numbers following the dash. An (E) after the page number indicates Erratum.

- Abe, Masayuki-(4) 1512
 Abernathy, C. R.-(1) 98; (4) 983
 Abraham, M.-(3) 579
 Adegboyega, G.-(3) 623
 Adesida, I.-(1) 49
 Agarwal, A.-(4) 1065
 Ahmed, H.-(1) 45; (2) 306; (4) 1402
 Ahmed, M. M.-(2) 306
 Ahn, C. C.-(4) 1108
 Akahori, T.-(1) 60
 Aketagawa, Masato-(3) 574
 Akinwande, A. I.-(2) 535
 Alimova, A. N.-(2) 446
 Anderson, S. G. H.-(4) 908
 Arnold, W.-(4) 1506
 Ascoli, Cesare-(4) 1398
 Ashok, S.-(2) 226
 Asp Möller Andersen, Bo-(4) 993
 Aspnes, D. E.-(4) 1196, 1205
 Azuma, T.-(2) 198
- Babin, Sergey-(2) 311
 Bachelet, F.-(2) 385
 Bai, C. L.-(4) 1343, 1350
 Bai, Chunli-(4) 1378
 Bai, Yubai-(4) 1442
 Baik, Young-Joon-(2) 499
 Bailey III, A. D.-(2) 373
 Bair, A. E.-(4) 1108
 Bandić, Z. Z.-(4) 1148
 Bandow, Shunji-(4) 1338
 Banerjee, A.-(4) 1074
 Banerjee, I.-(4) 955
 Baptist, R.-(2) 385
 Bar, I.-(4) 823
 Barratt, C.-(4) 983
 Baschieri, Paolo-(4) 1398
 Baskin, L. M.-(2) 410
 Baumann, P. K.-(4) 1236
 Baur, C.-(4) 1577
 Beard, B. T.-(1) 116
 Bedair, S. M.-(4) 1139
 Behm, R. J.-(4) 1373
 Behrendt, F.-(4) 1373
 Bell, F. H.-(1) 88; (3) 629
 Bell, K. A.-(4) 1196, 1205
 Beltram, Fabio-(4) 1398
 Berkó, A.-(1) 25
 Bernholc, J.-(4) 1144
 Berruyer, P.-(3) 629
 Berryman, K. W.-(4) 1045
 Besombes, C.-(4) 854
 Bessolov, Vasily N.-(4) 876
 Best, M. E.-(4) 1584
 Bezryadin, A.-(4) 793
 Bijkerk, F.-(2) 293
 Blanc, R.-(3) 629
 Blue, L. J.-(3) 696
 Bode, M.-(4) 1330
 Bodnar, S.-(3) 712
 Bogdanov, A. L.-(4) 814
 Bonard, J.-M.-(4) 1279
 Bonnell, Dawn A.-(4) 1483
 Booth, J. L.-(1) 116
 Bor, Zs.-(2) 287
 Borusik, O.-(4) 1445
 Bösch, M.-(4) 1564
 Botchkarev, A. E.-(4) 921
 Bouillon, P.-(3) 712
- Bourret, E. D.-(4) 891
 Bourret, Edith D.-(2) 247
 Boyd, R. D.-(3) 729
 Brake, J.-(3) 605, 670
 Bresse, J. F.-(4) 854
 Brillert, Ch.-(4) 1564
 Britten, J. A.-(3) 729
 Brown, W. L.-(4) 1065
 Brugat, Manuel-(2) 405
 Bruley, J.-(4) 840
 Buckley, Leonard J.-(2) 259; (3) 741
 Burke, T. M.-(2) 325
 Burroughes, J. H.-(2) 325
- Caffin, D.-(4) 854
 Camphausen, S. M.-(4) 840
 Cao, Yaan-(4) 1442
 Cao, Yunwei-(4) 1425
 Cappy, Alain-(4) 1008
 Caudano, Y.-(4) 1065
 Cavallaro, J. R.-(2) 287
 Cecconi, Ciro-(4) 1398
 Cha, C. Y.-(3) 605, 670
 Cháb, V.-(4) 1445
 Chabal, Y. J.-(4) 1065
 Chaban, E. E.-(4) 1065
 Chai, Xiangdong-(4) 1425
 Chakhovskoi, A. G.-(2) 507
 Chakhovskoi, Andrei G.-(2) 401, 516
 Chang, C. P.-(3) 646
 Chang-Liao, K. S.-(4) 942
 Chen, D. M.-(4) 1295
 Chen, H.-(4) 1456
 Chen, H. H.-(3) 681
 Chen, H. Y.-(4) 1581
 Chen, Jian-Guang-(3) 707
 Chen, L. C.-(4) 942
 Chen, Q.-(4) 1117
 Chen, S. F.-(4) 1573
 Chen, Sheng-fu-(4) 1385
 Chen, W.-(4) 1402
 Chen, X. L.-(4) 1502
 Chen, Y. J.-(4) 809
 Chen, Yongmei-(4) 1442
 Chen, Zhenghao-(4) 1112
 Chen, Zhi-(2) 252
 Cheng, H. K.-(4) 1498
 Cheng, K. Y.-(4) 1034
 Chernozatonskii, L. A.-(2) 422
 Cheung, W. Y.-(4) 809
 Chiang, Wen-Yen-(2) 299
 Chilla, E.-(4) 1569
 Choi, Duck-Kyun-(4) 928
 Choi, J. H.-(2) 528
 Choi, W. B.-(4) 840
 Choi, Woo-Beom-(2) 477
 Chow, D. H.-(2) 329
 Chowdhury, Ashfaqu I.-(1) 127
 Christman, S. B.-(4) 1065
 Chu, Jiaru-(4) 1551
 Chubun, N. N.-(2) 442, 450
 Chung, C.-H.-(4) 1182
 Chung, Hong-Bay-(4) 818
 Chung, Suk Jae-(2) 428, 431, 454
 Chvoj, Z.-(4) 1445
 Claflin, B.-(4) 1074
 Clark, S. A.-(3) 687
 Clement, M.-(4) 903
 Constancias, C.-(2) 385
- Constantine, C.-(4) 983
 Cornet, A.-(3) 687
 Corrêa F., J. A.-(4) 870
 Correia, A.-(3) 548
 Cox, J. N.-(4) 955
 Craighead, Harold G.-(2) 343
 Croke, E. T.-(4) 1108
 Cue, Nelson-(2) 273
 Cui, D. F.-(4) 1573
 Cunningham, J. E.-(4) 886
 Cuomo, J. J.-(2) 202; (4) 840
 Cutler, P. H.-(2) 337, 398
 Czuprynski, P.-(3) 629; (4) 1000
- Dachun, Zhao-(4) 805
 Dai, Liming-(3) 724
 Daniel, E. S.-(4) 1089
 Daniels, C.-(3) 702
 Daniels-Race, T.-(3) 696
 da Silva, M. I. N.-(4) 870
 Das Sarma, S.-(4) 1051
 Davidson, J. L.-(2) 460
 de Berranger, E.-(3) 712
 De Feyter, S.-(4) 1419
 de Groot, L. E. M.-(2) 293
 Dekker, C.-(3) 586; (4) 793
 DeLong, M. C.-(4) 1201
 de Lyon, T. J.-(2) 329
 Deng, Huihua-(4) 1460
 Denton, John P.-(1) 138
 de Oliveira, A. G.-(4) 870
 De Schryver, F. C.-(4) 1419
 Diebel, J.-(3) 579
 Diéguez, A.-(3) 687
 Ding, M. Q.-(4) 840
 Ding, Weiping-(4) 1347
 Dobisz, E.-(4) 1369
 Dong, G. O.-(4) 849
 Dong, H. K.-(1) 159
 Dorsey, D. L.-(4) 1159
 Dostalík, W. W.-(4) 967
 Downer, M. C.-(4) 1059
 Drews, D.-(4) 1128
 Drozdov, A. V.-(2) 482
 Du, Y. W.-(4) 1456
 Du, Youwei-(4) 1347, 1465
 Dubey, M.-(4) 990
 Duxstad, K. J.-(4) 891
 Duxstad, Kristin J.-(2) 247
- Eaglesham, D. J.-(4) 1065
 Eastman, L. F.-(4) 1187, 1274
 Eguchi, K.-(4) 1429
 Ehrfeld, W.-(3) 579
 Eisenberg, N. P.-(4) 823
 Ekerdt, J. G.-(4) 1059
 El-Gomati, M. M.-(2) 434
 Erdélyi, M.-(2) 287
 Esashi, M.-(4) 1414
 Esashi, Masayoshi-(4) 1531
 Escamilla, Hector-(4) 1516
 Esser, N.-(4) 1260
 Etoh, K.-(4) 788
 Evans, K. R.-(4) 1159
 Evtukh, A. A.-(2) 439
 Eyink, K. G.-(4) 1187, 1274
- Fang, Jinghuai-(4) 1468
 Farnault, E.-(4) 1539, 1556
 Fedotov, A. B.-(4) 1112
 Feenstra, R. M.-(4) 1027
 Felter, T. E.-(2) 507
 Ferincz, I. E.-(4) 828
 Fernandez, A.-(3) 729
 Ferry, D. K.-(4) 899
 Filippi, R. G.-(3) 750
 Flörsheimer, M.-(4) 1564
 Förster, A.-(3) 702
 Fourné, Hervé-(4) 1008
 Franciosi, A.-(4) 1279
 Franke, J. E.-(4) 955
 Frediani, Carlo-(4) 1398
 Freundt, D.-(4) 1121
 Friedrich, J.-(2) 293
 Frisch, A. M.-(4) 1260
 Fröhlich, H. J.-(4) 1569
 Fuchs, H.-(4) 1564
 Fujii, Toru-(4) 1494
 Fujii, Toshiaki-(3) 746
 Fujisawa, Satoru-(4) 1479
 Fujita, H.-(4) 1556
 Fujita, Ken-(1) 40
 Fukuda, Nobuhiro-(1) 53
 Fukuda, Shin-(1) 53
 Fukui, Takashi-(4) 1163
 Fursey, G. N.-(2) 410
 Furutake, Naoya-(2) 488
- Ganguli, Seshadri-(4) 948
 Ganière, J.-D.-(4) 1279
 Gao, H. J.-(4) 1581
 Gao, Juning-(4) 1313
 Gao, P. J.-(4) 1502
 García, N.-(3) 548
 Gastev, S. V.-(4) 1471
 Gayet, P.-(1) 133
 Gazen, B. C.-(4) 1577
 Geerligs, L. J.-(3) 586
 Geisz, J. F.-(4) 1201
 George, M. A.-(2) 460
 Gerthsen, D.-(4) 1121
 Giapis, Konstantinos P.-(1) 70
 Gibson, J. M.-(4) 921
 Giewont, Kenneth J.-(4) 916
 Gill, David M.-(3) 719
 Gill, William N.-(4) 948
 Givargizov, E. I.-(2) 442, 446, 450
 Glukhova, O. E.-(2) 422
 Göbel, H.-(4) 1359
 Gogolides, E.-(3) 640
 Goh, G. P.-(3) 610
 Goh, M. Cynthia-(2) 181
 Goldman, R. S.-(4) 1027
 Götz, A.-(4) 1097
 Goodman, Rory S.-(4) 971
 Gorfinkel, B. I.-(2) 528
 Gorfinkel, Boris-(2) 524
 Gorwadkar, Sucheta M.-(4) 1406
 Gosain, D. P.-(3) 554
 Goto, Toshio-(3) 746
 Gottscho, R. A.-(2) 373
 Gregus, J. A.-(2) 373
 Griffith, James-(3) 741
 Grigoriev, Y. A.-(2) 503
 Grigoropoulos, S.-(3) 640
 Grim, P. C. M.-(4) 1419

- Grimshaw, M. P.-(2) 325
 Gruber, A.-(4) 1382
 Grundbacher, R.-(1) 49
 Gspann, J.-(4) 1382
 Gu, Gang-(4) 1347
 Gu, Jianhua-(4) 1468
 Gu, Minming-(4) 1325
 Gu, Ning-(1) 178
 Guinn, K. V.-(2) 214
 Gulyaev, Yu. V.-(2) 391, 422, 533
 Guo, Zaibing-(4) 1347
 Gupalo, M. S.-(2) 491
 Güthner, P.-(3) 579
- Haak, M.-(4) 1521
 Haaland, D. M.-(4) 955
 Haas, T. W.-(4) 1187
 Hagmann, Mark J.-(2) 405; (3) 597
 Haller, E. E.-(2) 247; (4) 891
 Hallmark, J. A.-(4) 899
 Hamanaka, H.-(4) 1414
 Hamers, Robert J.-(4) 1153
 Han, B. Y.-(3) 605, 670
 Han, J. H.-(4) 1573
 Han, Min Koo-(2) 431
 Hankins, O. E.-(2) 202
 Hansen, Ole-(4) 993
 Hansson, G. V.-(2) 241
 Hao, H. W.-(4) 1498
 Haraichi, Satoshi-(4) 1406
 Harper, J. M. E.-(4) 763
 Hartmann, U.-(4) 1364
 Hasegawa, Hideki-(4) 1163, 1227
 Hasegawa, Y.-(4) 1270, 1300, 1307
 Hashizume, Tamotsu-(4) 1227
 Haskard, M. R.-(2) 477
 Hasko, D. G.-(1) 45
 Hasunuma, R.-(4) 1437
 Hattangady, S.-(4) 967
 Hattori, Shuzo-(3) 746
 Haworth, L.-(4) 1133
 Hawryluk, A. M.-(3) 729
 Hayasaki, Kei-(3) 558
 He, Huixin-(4) 1425
 He, Yuliang-(4) 1313
 Heerens, W. Chr.-(3) 566
 Heitzmann, M.-(4) 1000
 Herrmann, A.-(4) 1521
 Herrmann, T.-(4) 1260
 Hertz, H. M.-(4) 814
 Hesjedal, T.-(4) 1569
 Heun, S.-(4) 1279
 Hierlemann, M.-(4) 935
 Higuchi, Katsuhiko-(4) 862
 Higuchi, Toshiro-(1) 174; (4) 780
 Hill, P.-(4) 1133
 Hinds, B.-(4) 1074
 Hinode, K.-(2) 237
 Hinrichs, K.-(4) 1260
 Hiramatsu, Mineo-(3) 746
 Hiroshima, Hiroshi-(4) 1406
 Hirsekorn, S.-(4) 1506
 Ho, P. S.-(4) 908
 Hodel, U.-(4) 1364
 Holz, D.-(4) 1121
 Homma, Y.-(2) 237
 Hong, J.-(3) 652
 Hong, J. P.-(2) 528
 Hong, Qinyue-(1) 178
 Hong, Xiaoyin-(3) 724
 Hori, Masaru-(3) 746
 Horváth, Z. L.-(2) 287
 Hosaka, S.-(4) 788
 Hosch, Jimmy W.-(2) 361
 Houmady, M.-(4) 1539, 1556
 Hovis, Jennifer S.-(4) 1153
- Howard, A. J.-(3) 657
 Howe, Roger T.-(1) 1
 Hren, J. J.-(4) 840
 Hu, E. L.-(4) 1182
 Hu, Henry S.-(3) 741
 Hu, Jun-(4) 1385
 Hu, X. F.-(4) 1059
 Hua, Z. Y.-(4) 1353
 Huang, Feng-(4) 780
 Hudek, Peter-(2) 311
 Hughes, Padraig J.-(4) 845
 Hui, Zheng-(4) 1452
 Hulseweh, T.-(1) 122
 Hummler, Klaus M.-(4) 916
 Hunt, C. E.-(2) 507
 Hunt, Charles E.-(2) 401, 516
 Hunt, Douglas B.-(4) 916
 Hunter, A. T.-(4) 1108
 Hwang, Gyeong S.-(1) 70
 Hwang, W.-Y.-(2) 316
- Ibbotson, D. E.-(3) 646
 Ichikawa, Masakazu-(1) 40
 Iga, R.-(1) 103
 Igarashi, T.-(2) 277
 Ikoma, Y.-(2) 277
 Imai, K.-(1) 171
 Ingfield, C. E.-(4) 1201
 Irene, E. A.-(4) 1173
 Ishii, Kenichi-(3) 543; (4) 1406
 Ishikawa, T.-(1) 167
 Ishikawa, Yasuhiko-(4) 1163
 Ito, Yasuyuki-(4) 961
 Itoh, Toshihiro-(4) 1551, 1559
 Iwakuro, Hiroaki-(2) 232
 Iwamori, Satoru-(1) 53
 Izumi, S.-(1) 167
- Jackson, Howard E.-(4) 845
 Jacobs, L.-(4) 1359
 Jacobson, D. C.-(4) 1065
 Jalonen, Marko-(1) 154
 Jan, W. Y.-(4) 886
 Jang, J. E.-(2) 528
 Jang, Jin-(2) 428, 431, 454
 Jansons, A.-(4) 880
 Jelenkovic, Emil V.-(2) 209
 Jeon, Byung Soo-(2) 520
 Jeon, D.-(2) 425, 457
 Jeon, Min-Seok-(4) 928
 Jiang, H. X.-(4) 1117, 1139
 Jiang, X. B.-(4) 849
 Jiang, Yueshan-(4) 1313
 Jiang, Yueshun-(4) 1425
 Jiménez-Leube, F. J.-(4) 903
 Jin, C. G.-(4) 1212
 Jin, Haiyan-(4) 1442
 Jing, Xiao Gong-(4) 1452
 Johnson, B. R.-(2) 535
 Johnson, S. R.-(1) 116
 Jones, J. T.-(4) 1089
 Joubert, O.-(1) 88; (3) 629; (4) 1000
 Ju, Byeong Kwon-(2) 425, 431, 499, 512
 Ju, Byeong-Kwon-(2) 477
 Juan, W. H.-(2) 267; (3) 665
 Jung, Jae Hoon-(2) 431
- Kadota, Y.-(1) 103
 Kamal, A. H. M.-(4) 899
 Kamoshida, Kazuyoshi-(4) 961
 Kampen, T. U.-(4) 1221
 Kanao, H.-(1) 60
 Kandidov, A. V.-(2) 446
 Kang, J. H.-(2) 528
 Kang, Sang Won-(2) 520
- Kang, W. P.-(2) 460
 Kania, D. R.-(3) 729
 Kao, Y.-C.-(1) 49
 Kasai, Jun-ichi-(4) 862
 Katayama, Yoshifumi-(4) 862
 Kato, Tadao-(3) 614
 Kaupp, G.-(4) 1521
 Kawai, Tomoji-(3) 602
 Kawakatsu, H.-(4) 1539, 1556
 Kawasaki, H.-(4) 908
 Kendall, R. E.-(3) 696
 Kerns, D. V.-(2) 460
 Ketterson, A. A.-(1) 49
 Khan, M. A.-(4) 1117
 Khor, K. E.-(4) 1051
 Kiely, James D.-(4) 1483
 Kikukawa, A.-(4) 788
 Kilpelä, O.-(4) 1471
 Kim, Hwan-(2) 468
 Kim, J.-(4) 1011
 Kim, J. M.-(2) 528
 Kim, J. W.-(2) 528
 Kim, Jong Min-(2) 524
 Kim, Seongjin-(2) 499
 Kim, Wook-(4) 921
 Kim, Yeo Hwan-(2) 468
 Kimizuka, Masakatsu-(1) 66; (2) 221
 Kimura, K.-(4) 1212, 1241, 1254, 1254
 Kinsky, J.-(4) 1260
 Kishi, K.-(1) 103
 Kitano, T.-(1) 167
 Kiyama, H.-(4) 1270
 Klebanov, M.-(4) 823
 Kleiber, M.-(4) 1330
 Klemens, F. P.-(3) 646
 Klyachko, D.-(4) 1295
 Knipp, P. A.-(4) 1040
 Kobayashi, Katsuyoshi-(4) 1479
 Kochhar, R.-(2) 316
 Kodate, J.-(1) 171
 Koel, B.-(4) 1577
 Koh, K.-(4) 1074, 1097
 Koh, Michael T. K.-(1) 138
 Komeda, T.-(4) 1437
 Komiyama, Masaharu-(4) 1325
 Kondo, S.-(2) 237
 Konenkova, Elena V.-(4) 876
 Kontorovich, E. L.-(2) 495
 Koops, H. W. P.-(4) 1369, 1535
 Kornblit, A.-(3) 646
 Kosakovskaya, Z. Ya.-(2) 422
 Koster, N. B.-(2) 293
 Köster, P.-(4) 1527
 Kostic, Ivan-(2) 311
 Koyama, Hiroshi-(3) 614
 Koyanagi, H.-(4) 788
 Kozen, A.-(3) 618
 Kozicki, M. N.-(4) 899
 Kraft, R.-(4) 967
 Kravitz, S. H.-(3) 657
 Kristensen, Martin-(4) 993
 Kriván, E.-(1) 25
 Krolkowski, W.-(1) 122
 Kudryashov, V. P.-(2) 503
 Kuo, L. H.-(4) 1212, 1241, 1254
 Kuroda, R.-(4) 1429
 Kuroda, Tsukasa-(2) 232
 Kurz, H.-(4) 1097
 Kusumi, Yukihiro-(1) 40
 Kuyanov, I. A.-(1) 21
 Kvam, Eric P.-(1) 138
 Kwan, Siu L.-(1) 116
 Kwon, Sang Jik-(2) 472
 Kydzinovski, S. Yu.-(2) 439
- LaBella, V. P.-(4) 1191
 Lacher, M.-(3) 579
 Lajoie, E.-(4) 1000
 Lambers, E. S.-(3) 652
 Lampert, W. V.-(4) 1187
 Landgren, G.-(4) 1027
 Launay, P.-(4) 854
 Launois, H.-(1) 142
 Laursen, T.-(4) 1108
 Lavoie, C.-(4) 1265
 Lawther, D. W.-(3) 757
 Layadi, Nacer-(1) 109
 Lazzarino, Marco-(4) 1398
 Lebedev, Mikhail V.-(4) 876
 Lee, Chengkuo-(4) 1551, 1559
 Lee, Chun Gyo-(2) 464, 468
 Lee, Hyun-Yong-(4) 818
 Lee, I.-M.-(4) 880
 Lee, I-Ming-(1) 138
 Lee, J. W.-(1) 98; (3) 652; (4) 983
 Lee, John Tseng-Chung-(1) 109
 Lee, Jong Duk-(2) 464, 468, 472, 520
 Lee, Kwang-Ryeol-(2) 457
 Lee, Sanjo-(2) 425, 457
 Lee, Seung-(4) 1153
 Lee, Si-Chen-(2) 321
 Lee, Sungwoon-(2) 457
 Lee, Sunnup-(2) 457
 Lee, T. C.-(3) 646
 Lee, Yun Hi-(2) 425, 499, 512
 Lee, Yun-Hi-(2) 477
 Lefebvre, P. R.-(4) 1173
 Legay, P.-(4) 854
 Leone, Stephen R.-(4) 971
 Lerner, P.-(2) 398
 Lerner, Peter-(2) 337
 Leung, Edward C. W.-(2) 181
 Levine, E. N.-(3) 750
 Li, J. Z.-(4) 1117
 Li, L.-(4) 1300, 1307
 Li, Lin Song-(4) 1452
 Li, Min-qian-(4) 1385
 Li, N. Y.-(1) 159
 Li, Q.-(2) 460
 Li, Tie Jin-(4) 1452
 Li, Tiejun-(4) 1425, 1442
 Li, W. G.-(4) 1350
 Li, Xiaodong-(4) 1378
 Liang, Bingjie-(4) 1432
 Liliental-Weber, Z.-(4) 891
 Liliental-Weber, Zuzanna-(2) 247
 Lim, D.-(4) 1059
 Lim, Sungkyoo-(2) 499
 Lin, Hao-Hsiung-(3) 707
 Lin, J. Y.-(4) 1117, 1139
 Lin, Ray-Ming-(2) 321
 Lin, X. W.-(4) 891
 Lin, Zhangda-(2) 232
 Litovchenko, V. G.-(2) 439
 Liu, Bing-Der-(2) 321
 Liu, Hongbing-(4) 1153
 Liu, J.-(4) 1502, 1573
 Liu, J. H.-(4) 849
 Liu, Jin-Shung-(3) 707
 Liu, Ning-(4) 1313
 Liu, Yong Kuan-(1) 178
 Liu, Zhongfan-(4) 1425
 Logan, J. A.-(4) 1584
 Loginov, M. V.-(2) 495
 Löhndorf, M.-(4) 1330
 Lord, S. M.-(2) 316
 Louis, D.-(4) 1000
 Louis, E.-(2) 293
 Lovejoy, M. L.-(3) 657
 Loy, M. M. T.-(4) 1112
 Lu, J.-(4) 1133, 1456
 Lu, Jianping-(3) 724

- Lu, Ran-(4) 1425
 Lu, Xiaomei-(4) 1468
 Lu, Zuhong-(1) 178; (4) 1460, 1468
 Lucovsky, G.-(4) 1074, 1097
 Ludeke, R.-(4) 1080
 Luo, K.-(2) 349
 Lüth, H.-(3) 702; (4) 1121
 Lützen, J.-(4) 899
 Lynch, K. E.-(4) 990
 Lyon, S. A.-(4) 1045
 Lyubin, V.-(4) 823
- Ma, L. P.-(4) 1581
 Maboudian, Roya-(1) 1
 Macdonald, J. E.-(4) 1133
 Machida, K.-(1) 171
 Machorro, Roberto-(4) 1516
 MacKenzie, J. D.-(1) 98
 Maeda, Ryutaro-(4) 1559
 Mahalingam, K.-(4) 1159
 Majumdar, A.-(2) 349
 Maki, Kunisuke-(2) 192
 Malinowski, M. E.-(2) 507
 Malmqvist, L.-(4) 814
 Mamin, H. J.-(4) 1584
 Manassen, Y.-(4) 1310, 1317
 Manevich, M.-(4) 823
 Mantese, L.-(4) 1196, 1205
 Mao, Haifang-(4) 1460
 Mao, Haitao-(4) 1465
 Marchenko, R. I.-(2) 439
 Markiewicz, Peter-(2) 181
 Marqués, M. I.-(3) 548
 Marsh, O. J.-(4) 1089
 Marsico, V. E.-(4) 1065
 Masuzawa, T.-(4) 1556
 Materer, N.-(4) 971
 Matsuda, H.-(4) 1429
 Matsumoto, S.-(1) 103
 Matsuo, Seitaro-(3) 736
 Mau, Albert W.-(3) 724
 Maury, D.-(1) 133
 Mawatari, H.-(1) 103
 Mayer, J. W.-(4) 1108
 Mayer, T. S.-(2) 316
 Maynard, H. L.-(3) 646
 Maynard, Helen L.-(1) 109
 Mayr, K.-(3) 579
 McCaldin, J. O.-(4) 1148
 McGill, T. C.-(4) 1089, 1148
 McIntosh, F. G.-(4) 1139
 McLane, G. F.-(4) 990
 McMurray, J. S.-(4) 1011
 McNeil, John R.-(2) 361
 McNevin, S. C.-(2) 214
 Memmert, U.-(4) 1364
 Merli, M.-(3) 623
 Mevlyut, Sh. T.-(2) 422
 Micovic, M.-(2) 316
 Miller, D. L.-(2) 316
 Milne, W. I.-(2) 428, 431
 Milosavljevic, I.-(2) 460
 Minami, H.-(1) 167
 Minggen, Sun-(4) 805
 Miskovsky, N.-(2) 398
 Miskovsky, N. M.-(2) 337
 Mitchel, W. C.-(4) 1187
 Mitchell, W. J.-(4) 1182
 Miwa, S.-(4) 1212, 1241, 1254
 Miyamoto, M.-(4) 788
 Miyano, Soichiro-(2) 488
 Miyashita, Masayuki-(4) 1547
 Miyashita, Takehiro-(1) 53
 Miyazaki, H.-(2) 237
 Mizutani, W.-(4) 1388
 Moessner, G.-(4) 1419
- Mohammad, S. N.-(4) 921
 Mohammad, S. Noor-(2) 252
 Mönch, W.-(4) 1221
 Montelius, L.-(4) 814
 Moon, Jong Hyun-(2) 428, 431, 454
 Morante, J. R.-(3) 687
 Morita, Seizo-(4) 1479, 1512, 1543
 Morkoç, H.-(4) 921
 Morkoç, Hadis-(2) 252
 Motomatsu, M.-(4) 1388
 Motooka, T.-(2) 277
 Mouis, M.-(3) 712
 Moulin, A. M.-(3) 590
 Moy, A. M.-(4) 1034
 Muji, Dai-(4) 805
 Mukaida, H.-(4) 1437
 Müllen, K.-(4) 1419
 Murashita, T.-(1) 32
 Murnane, Michael R.-(2) 361
 Murphy, D.-(2) 535
 Myers, A. F.-(4) 840
 Myler, U.-(3) 757
 Myrick, M. L.-(2) 282
- Nakagiri, Nobuyuki-(4) 1394, 1547
 Nakagiri, T.-(4) 1429
 Nakajima, K.-(4) 1254
 Nakamura, K.-(4) 788
 Nakashima, Hiroshi-(4) 1411
 Nakayama, Kan-(4) 1494
 Nam, C. W.-(2) 226
 Naqvi, H.-(2) 361
 Narayan, J.-(4) 840
 Nardelli, Marco Buongiorno-(4) 1144
 Nassiopoulou, A. G.-(3) 640
 Nawata, Masahito-(3) 746
 Nemanich, R. J.-(4) 1236
 Neudeck, Gerold W.-(1) 138
 Ng, K. Y. S.-(2) 186
 Nguyen, H. T.-(3) 729
 Ni, G.-(4) 1456
 Nie, H. Y.-(4) 1388
 Niemczyk, T. M.-(4) 955
 Nieto, Jesús-(4) 1516
 Niimi, H.-(4) 1097
 Nishioka, Tadashi-(3) 614
 Nishioki, Nobuhisa-(1) 174
 Nisimura, Miyo-(2) 488
 Nissim, Y.-(1) 142
 Noll, J. D.-(2) 282
 Novikov, S. V.-(4) 1471
 Nur, O.-(2) 241
- Ogino, T.-(4) 1270
 Oh, Myung Hwan-(2) 425, 428, 431, 499, 512
 Oh, Myung-Hwan-(2) 477
 Ohashi, Takahiro-(4) 1559
 Ohler, C.-(3) 702
 Ohshima, Kazunari-(4) 1338
 Ohta, Masahiro-(4) 1512, 1543
 Ohtake, A.-(4) 1212, 1241, 1254
 Okada, Ikuo-(3) 736
 Okada, K.-(4) 1449
 Okamoto, Akihiko-(2) 488
 Okamoto, H.-(1) 103
 Olami, Z.-(4) 1317
 Olson, J. M.-(4) 1201
 Ono, T.-(4) 1414
 Ono, Takahito-(4) 1531
 Ono, Toshiro-(3) 736
 Otsubo, M.-(1) 167
 Ouyang, M.-(4) 1304
 Ouyang, Q.-(4) 1449
 Ouyang, Zhen-qian-(4) 1385
 Owens, D. W.-(3) 605, 670
- Ozaki, Yoshiharu-(1) 66; (2) 221
- Paggel, J. J.-(4) 1279
 Pai, Woei Wu-(4) 785
 Pang, S. J.-(4) 1502, 1581
 Pang, S. W.-(2) 267; (3) 665, 681
 Pang, Shijin-(4) 1313
 Park, Beom Soo-(2) 499
 Park, Byung Gook-(2) 464, 468
 Park, Dae-Gyu-(2) 252
 Park, Kyu Chang-(2) 428, 431, 454
 Park, N. S.-(2) 528
 Parsons, Gregory N.-(1) 127
 Pascal, R.-(4) 1330
 Patriarche, G.-(4) 854
 Pearsall, T. P.-(1) 116
 Pearton, S. J.-(1) 98; (3) 652; (4) 983
 Peng, Li-(1) 178
 Penzyakov, V. V.-(2) 503
 Pépin, A.-(1) 142
 Pepper, M.-(2) 325
 Perez-Quintana, I.-(3) 623
 Perry, M. D.-(3) 729
 Pesant, Jean Claude-(4) 1008
 Petrosyan, A. I.-(2) 503
 Pimenov, V. G.-(2) 503
 Piner, E. L.-(4) 1139
 Pingue, Pasqualantonio-(4) 1398
 Pinnington, T.-(4) 1265
 Piquette, E. C.-(4) 1148
 Pistecsky, P. V.-(3) 566
 Poggi, A.-(3) 623
 Poon, M. C.-(2) 394
 Posadowski, W. P.-(2) 202
 Prins, Steven L.-(2) 361
 Pyle, Jason L.-(1) 38
- Qian, Feng-(1) 178
 Qiang, D.-(4) 1304
 Qiao, H.-(4) 1498
 Qiu, Xiaohui-(4) 1378
 Quierin, M. A.-(2) 325
- Rabe, U.-(4) 1506
 Rack, M. J.-(4) 899
 Radamson, H. H.-(2) 241
 Radzinski, Z. J.-(2) 202
 Ramachandran, T. R.-(4) 1577
 Ramesh Rao, N.-(4) 1317
 Rapcewicz, Krzysztof-(4) 1144
 Rauscher, H.-(4) 1373
 Ray, Mark-(3) 741
 Raymond, Christopher J.-(2) 361
 Read, Walter W.-(1) 127
 Regolini, J. L.-(1) 133; (3) 712
 Reinecke, T. L.-(4) 1040
 Reinhardt, J.-(3) 579
 Ren, F.-(4) 983
 Requicha, A. A. G.-(4) 1577
 Ribeiro, G. M.-(4) 870
 Rich, D. H.-(4) 1034
 Richter, W.-(4) 1260
 Rieger, D. J.-(3) 657
 Rishton, S. A.-(4) 1584
 Ritchie, D. A.-(2) 325
 Rizzi, A.-(4) 1121
 Robert, J. C.-(4) 1139
 Rodbell, K. P.-(3) 750; (4) 763
 Rodríguez, T.-(4) 903
 Rogovin, V. I.-(2) 503
 Roman, M.-(4) 1121
 Romijn, J.-(2) 293
 Rosenwaks, S.-(4) 823
 Rossow, U.-(4) 1196, 1205, 1260
 Rostoll, M. L.-(1) 133
 Roth, J. A.-(2) 329
- Rousseau, P. M.-(3) 757
 Rousseau, B. A. C.-(2) 293
 Roussina, E. V.-(2) 528
 Roux, G. Le-(4) 854
 Rubini, S.-(4) 1279
 Rubloff, Gary W.-(1) 127
 Rücker, M.-(4) 1419
 Ruda, H.-(3) 554
 Ruf, A.-(3) 579
 Rugar, D.-(4) 1584
 Ruskell, Todd G.-(1) 38
 Ruvimov, S.-(4) 891
 Ruvimov, Sergei-(2) 247
 Ryu, Y. S.-(2) 528
- Saitoh, Hiroaki-(4) 1531
 Sakai, Masahiro-(4) 1338
 Sako, Sanshiro-(4) 1338
 Sakuma, N.-(2) 237
 Sakurai, T.-(4) 1270, 1300, 1307
 Salokatve, Arto-(1) 154
 Sampaio, J. F.-(4) 870
 Sang, H.-(4) 1456
 Sanz Maudes, J.-(4) 903
 Sappjeta, J.-(4) 1065
 Sardela, M. R., Jr.-(2) 241
 Sarid, Dror-(1) 38
 Sasaki, Naruo-(4) 1479
 Sasaki, Shigetoshi-(3) 574
 Sato, K.-(1) 167
 Sato, Shinji-(2) 232
 Sato, T.-(1) 45; (2) 198
 Sato, Taketomo-(4) 1227
 Scherer, V.-(4) 1506
 Schmid, C. R.-(3) 696
 Schmitsdorf, R. F.-(4) 1221
 Schneider, A.-(4) 1128
 Schneider, M.-(1) 142
 Schneider, T. P.-(4) 967
 Schössler, C.-(4) 1535
 Schowalter, L. J.-(4) 1191
 Schuler, Olivier-(4) 1008
 Schultz, Ch.-(4) 1260
 Schwarz, U. D.-(4) 1527
 Seaford, M. L.-(4) 1187, 1274
 Segev, Mordechai-(4) 1045
 Sekigawa, Toshihiro-(3) 543
 Sekiguchi, T.-(2) 226
 Seleznev, B. V.-(2) 446
 Semyonov, V. C.-(2) 503
 Seyama, Hideo-(2) 192
 Shachal, D.-(4) 1310
 Shang, G. Y.-(4) 1343
 Shang, Guangyi-(4) 1378
 Shann, Hwai-Der-(2) 299
 Sheng, T. T.-(3) 610
 Shesterkin, V. I.-(2) 503
 Shi, Bo-Rong-(2) 273
 Shi, D. X.-(4) 1502
 Shi, Dongxia-(4) 1313
 Shi, Z.-(2) 349
 Shibuki, S.-(1) 60
 Shido, S.-(4) 1429
 Shimada, Masaru-(3) 736
 Shin, Dong Ki-(2) 512
 Shingubara, S.-(2) 202
 Shinohara, H.-(4) 1300
 Shirochin, L. A.-(2) 410
 Shrednik, V. N.-(2) 495
 Shul, R. J.-(1) 98; (3) 657; (4) 983
 Sigematsu, Tatsuhiko-(4) 1543
 Silfvenius, C.-(4) 1027
 Simpson, P. J.-(3) 757
 Sin, Johnny K. O.-(2) 394
 Sinitsyn, N. I.-(2) 422, 533
 Sinitsyn, N. L.-(2) 391

- Sinkkonen, J.-(4) 1471
 Siqueiros, Jesús-(4) 1516
 Siqueiros, Jesús M.-(4) 1474
 Skotnicki, T.-(3) 712
 Smayling, M. C.-(2) 287
 Smith, M.-(4) 1139
 Smith, Terry L.-(2) 273
 Snow, Arthur W.-(2) 259; (3) 741
 Sohail, S.-(2) 361
 Sohma, Yasunari-(2) 343
 Solomon, J. S.-(4) 1187
 Song, Man Ho-(2) 512
 Sonoda, T.-(1) 167
 Sorba, L.-(4) 1279
 Spitzer, A.-(4) 935
 Stålnacke, B.-(4) 1027
 Stein, B. L.-(4) 1108
 Stepanova, A. N.-(2) 450
 Stephenson, R. J.-(3) 590
 Suchorski, Yu.-(2) 491
 Sudoh, Kazufuyu-(1) 53
 Suga, Tatamoto-(4) 1551, 1559
 Sugawara, Yasuhiro-(4) 1479, 1512, 1543
 Sugimura, Hiroyuki-(4) 1394
 Sugimuraand, Hiroyuki-(4) 1547
 Sun, Jie-lin-(4) 1385
 Sun, Tai-Ping-(2) 321
 Sun, Yue-(4) 1465
 Sung, Man-Young-(2) 477
 Susi, E.-(3) 623
 Suzuki, Eiichi-(3) 543
 Suzuki, Masatoshi-(4) 1494
 Suzuki, Mineharu-(4) 1543
 Suzuki, Seiki-(3) 574
 Suzuki, Yoshihiko-(4) 1547
 Svensson, Stefan P.-(3) 719
 Swider, Wendy-(2) 247
 Szabó, G.-(2) 287
- Tabara, Suguru-(3) 675
 Tabet, M. F.-(4) 800
 Takada, Koji-(3) 574
 Takahashi, Hideaki-(3) 574
 Takamura, Yuzuru-(3) 558
 Takeda, K.-(2) 237
 Takemura, Hisashi-(2) 488
 Takimoto, K.-(4) 1429
 Takoudis, C. G.-(4) 880
 Takoudis, Christos G.-(1) 138
 Talin, A. A.-(2) 507
 Tanaka, Hiroyuki-(3) 602
 Tanaka, K.-(4) 1212
 Tanaka, M.-(1) 171
 Tanaka, Sigehisa-(4) 862
 Tanenbaum, David M.-(2) 343
 Tang, Y.-(4) 1034
 Tans, S. J.-(3) 586
 Taylor, J. Ashley-(2) 214
 Taylor, P. C.-(4) 1201
 Tedder, Laura L.-(1) 127
 Teitsworth, S. W.-(3) 696
 Terashima, Kazuo-(3) 558
 Ter Ovanessian, E.-(4) 1317
 Terris, B. D.-(4) 1584
 Thomas, S., III-(3) 681
 Tian, F.-(4) 1343, 1350
- Tiedje, T.-(4) 1265
 Tittel, F. K.-(2) 287
 Tlali, Spirit-(4) 845
 Tokumoto, H.-(4) 1388, 1437
 Tomich, D. H.-(4) 1187, 1274
 Tomiya, S.-(3) 554
 Tong, K. Y.-(2) 209
 Torgashov, G. V.-(2) 422
 Torgashov, I. G.-(2) 422
 Tóth, Cs.-(4) 828
 Townner, Frederick J.-(3) 719
 Trujillo, Johann T.-(2) 401
 Tserepi, A. D.-(3) 640
 Tsuda, Shunji-(2) 488
 Tsukada, Masaru-(4) 1479
 Tu, C. W.-(1) 159
 Tu, Yuan-Kuang-(3) 707
 Tucker, J. R.-(4) 1274
 Tung, C. H.-(3) 610
 Turan, R.-(2) 241
 Turner, Stephen W.-(2) 343
- Uchihashi, Takayuki-(4) 1512, 1543
 Uematsu, Y.-(2) 277
 Ueyama, Hitoshi-(4) 1512
 Uh, Hyung Soo-(2) 472
 Uozumi, Kiyohiko-(4) 1411
 Urban, F. K., III-(4) 800
 Urban, J.-(4) 1369, 1535
 Usui, S.-(3) 554
- Valiyaveetil, S.-(4) 1419
 van der Drift, E. W. J. M.-(2) 293
 van der Wulp, H.-(3) 566
 van de Sanden, M. C. M.-(2) 373
 Vanoppen, P.-(4) 1419
 Van Patten, P. G.-(2) 282
 Varesi, J.-(2) 349
 Vartuli, C. B.-(1) 98
 Venkat, Rama-(4) 1159
 Ventrice, C. A., Jr.-(4) 1191
 Vieu, C.-(1) 142
 Vilà, A.-(3) 687
 Vispute, R. D.-(4) 840
 Vizioz, C.-(4) 1000
 Vladimirov, G. G.-(2) 482
 von Blanckenhagen, P.-(4) 1359, 1382
 Voorma, H. J.-(2) 293
 Voronin, A. B.-(2) 442
- Wada, Toshimi-(4) 1406
 Wadas, A.-(4) 1330
 Walck, S. D.-(4) 1187
 Walck, S. N.-(4) 1040
 Wan, M. X.-(4) 1350
 Wang, B.-(4) 1449
 Wang, Baoping-(2) 394
 Wang, C.-(4) 1343, 1350
 Wang, Chen-(2) 394; (4) 1378
 Wang, K. Z.-(4) 1304
 Wang, L.-(4) 1573
 Wang, L. F.-(3) 610
 Wang, Li-(4) 833
 Wang, Li Jun-(4) 1452
 Wang, N. X.-(4) 1343
- Wang, Shu-(4) 1474
 Wang, Wei-Chung-(1) 138
 Wang, Yongqiang-(4) 1425
 Wang, Zhongze-(1) 116
 Warmack, R. J.-(4) 785
 Washburn, J.-(4) 891
 Washburn, Jack-(2) 247
 Watanabe, F.-(2) 277
 Watanabe, Kazutoshi-(4) 1551
 Watanabe, Yoshio-(1) 66
 Weaver, J. H.-(3) 605, 670
 Wegner, G.-(3) 586
 Wei, W. W.-(4) 1498
 Wei, Yu-(1) 178; (4) 1432
 Weigold, J. W.-(2) 267
 Weinberg, W. H.-(4) 1182
 Weiss, Bernard L.-(4) 845
 Weiss, D.-(4) 1330
 Weldon, M. K.-(4) 1065
 Welland, M. E.-(3) 590
 Wells, T.-(2) 434
 Wen, H. J.-(4) 1080
 Wendelken, John F.-(4) 785
 Werner, C.-(4) 935
 Wessels, B. W.-(4) 1056
 Westwater, J.-(3) 554
 Westwood, D. I.-(3) 687; (4) 1133
 Wierschem, M.-(4) 1564
 Wiesendanger, R.-(4) 1330, 1527
 Willander, M.-(2) 241
 Williams, C. C.-(4) 1011
 Wilson, I. H.-(4) 809
 Wilson, R.-(1) 122
 Wolfe, D.-(4) 1097
 Wong, S. P.-(4) 809
 Wong, W. S.-(1) 159
 Wood, J.-(2) 434
 Wood, M. C.-(4) 990
 Word, J. C.-(3) 657
 Workman, Richard K.-(1) 38
 Wu, H. Z.-(4) 849
 Wu, J.-(3) 586
 Wu, J. Z.-(4) 849
 Wu, Jingwen-(4) 1468
 Wu, Q. D.-(4) 1304
 Wu, W.-(4) 1274
- Xia, S. H.-(4) 1573
 Xiao, Mufei-(4) 1474, 1516
 Xiao, Xudong-(4) 1112
 Xiaoren, Pan-(4) 805
 Xie, Xide-(4) 1105
 Xu, H.-(2) 186
 Xu, Hongqi-(4) 1335
 Xu, Huijun-(4) 1460
 Xu, J. B.-(4) 809
 Xu, J. F.-(2) 460
 Xu, Jianfeng-(4) 1465
 Xu, Tian-Bing-(2) 273
 Xu, W.-(4) 1353
 Xu, Z.-(4) 1059
 Xue, Q. K.-(4) 1270
 Xue, Z. Q.-(4) 1304, 1581
 Xue, Zengquan-(4) 1313
 Yabumoto, N.-(1) 171
 Yahiro, T.-(4) 1539
 Yamada, Hirofumi-(4) 1494
- Yamada, Toshishige-(4) 1019
 Yamaguchi, K.-(2) 277
 Yamaguchi, Masataka-(4) 1494
 Yamamoto, N.-(1) 103
 Yamamoto, Takuma-(4) 1547
 Yamazaki, Nobuhiro-(2) 192
 Yanase, Yoshio-(4) 1543
 Yang, H.-(4) 1074
 Yang, H. C.-(2) 528
 Yang, Haiqiang-(4) 1313
 Yang, Jihua-(4) 1442
 Yang, Wensheng-(4) 1425, 1442
 Yang, Yongyuan-(3) 724
 Yao, T.-(4) 1212, 1241, 1254
 Yao, Xiaowei-(1) 38
 Yarish, I. L.-(2) 491
 Yasuda, S.-(4) 1429
 Yasuda, T.-(4) 1212, 1241, 1254
 Yasue, Takao-(3) 614
 Ye, Z. Z.-(4) 849
 Yeadon, M.-(4) 921
 Yeo, I. S.-(4) 908
 Yi, S. I.-(4) 1182
 Yi, W. K.-(4) 1498
 Ying, F.-(3) 665
 Yokoi, Tsunekio-(3) 746
 Yoo, Jae Soo-(2) 520
 Yoshida, Toyonobu-(3) 558
 Yoshiki, Masayuki-(2) 488
 Young, J. F.-(4) 828
 Yu, E. T.-(4) 1108
 Yu, K. M.-(4) 891
 Yu, Xiangdong-(1) 178
 Yuan, Chunwei-(4) 1432
 Yuda, M.-(3) 618
- Zahn, D. R. T.-(4) 1128
 Zakharchenko, Yu. F.-(2) 391, 422, 533
 Zavada, J.-(4) 1139
 Zavadinsky, V. G.-(1) 21
 Zeng, K. C.-(4) 1139
 Zhang, Chun-(4) 1112
 Zhang, H. X.-(4) 1581
 Zhang, Haijun-(1) 174; (4) 780
 Zhang, Haiping-(1) 178
 Zhang, J. R.-(4) 1456
 Zhang, Jiandi-(4) 785
 Zhang, Kaiming-(4) 1105
 Zhang, Ning-(4) 1347
 Zhang, S.-(4) 955
 Zhang, Y. Z.-(4) 1502
 Zhang, Zhongyi-(1) 178
 Zhao, Hongwei-(2) 394
 Zhao, Lixing-(1) 178
 Zhao, Yong Fang-(4) 1452
 Zhabanov, A. I.-(2) 422
 Zhirnov, V. V.-(2) 442, 446, 450
 Zhong, Wei-(4) 1347
 Zhongkai, Qu-(4) 805
 Zi, Jian-(4) 1105
 Zijlstra, T.-(2) 293
 Zimmerman, Neil M.-(2) 369
 Zini, L.-(4) 1577
 Zlupko, V. M.-(2) 491
 Zwörner, O.-(4) 1527

The American Vacuum Society Is ...



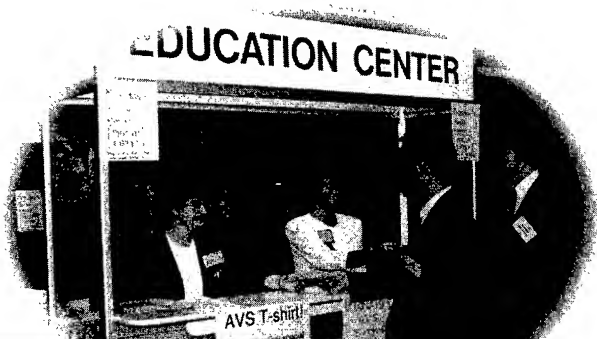
Partnering with industry.

The exhibition, held in conjunction with the AVS' annual symposium, is the largest display of the latest innovations in vacuum and deposition equipment, analysis systems, and related services.



Advancing technology worldwide.

Scientists and engineers from universities, government labs, research institutes, and industry gather each fall to attend the week-long AVS National Symposium and learn about the latest developments in their field and exchange ideas with professionals from around the world.



Educating the science community.

Our Educational Center at the annual symposium offers a wide variety of cost-effective educational resources covering vacuum and related technologies.



Providing practical job training.

AVS short courses train more than 3,000 engineers, managers, technicians, and others annually in vacuum technology, surface science, materials and materials characterization, and thin films and processing.



Recognizing outstanding achievements.

AVS awards encourage excellence in graduate students, providing grants for further research, and honor the scientific pioneers among AVS members.



For more information, contact the AVS, 120 Wall Street, 32nd Floor, New York, NY 10005, 212-248-0200, fax 212-248-0245, e-mail avsny@vacuum.org, <http://www.vacuum.org>.



AMERICAN VACUUM SOCIETY

Under United States copyright law, a transfer of copyright from the author(s) must be explicitly stated in writing to enable the publisher to disseminate the work to the fullest extent. The following transfer agreement must be completed, signed and returned to the Editor's office before the manuscript can be accepted for publication. Further information is available from the Office of Rights and Permissions, American Institute of Physics (AIP), 500 Sunnyside Boulevard, Woodbury, NY 11797-2999; Tel: (516) 576-2268; Fax: (516) 576-2327; Internet: rights@aip.org.

TRANSFER OF COPYRIGHT AGREEMENT

Copyright in the unpublished and original article, including the abstract forming part thereof, entitled _____

Title of Article

(the "Article"), submitted by the following author(s) [the "Author(s)"] _____

Names of all Authors

is hereby assigned and transferred to the American Vacuum Society ("AVS") for the full term thereof throughout the world, subject to the terms of this Agreement and to acceptance of the Article for publication in _____

Name of Journal

AVS shall have the right to publish the Article in any medium or form, or by any means, now known or later developed. AVS shall have the right to register copyright to the Article in its name as claimant whether separately or as part of the journal issue or other medium in which the Article is included. The Author(s) reserve all proprietary rights other than copyright, such as patent rights. If the Article was prepared under a United States Government contract, the Government shall have rights in the copyright to the extent required by the contract, and the Author(s) shall notify AVS of any such rights. The Author(s) represent and warrant: (1) that the Article is original with them; (2) that the Article does not infringe any copyright or other rights in any other work, or violate any other rights; (3) that the Author(s) own the copyright in the Article or are authorized to transfer it; and (4) that all copies of the Article the Author(s) make or authorize will include a proper notice of copyright in AVS's name. If each Author(s) signature does not appear below, the signing Author(s) represent that they sign this Agreement as authorized agents for and on behalf of all the Authors, and that this Agreement and authorization is made on behalf of all the Authors.

Name (print)

Date

Signature

If the manuscript has been prepared as a work for hire, this Agreement should be signed by the employee (or commissioned party) (**above**) and by the employer (or commissioning party) (**below**). In such event, the employer (or commissioning party) agrees to be bound by all of the obligations of the Author(s) under this Agreement.

Name of Employer (or Commissioning Party) (print)

Signature

Title

Date

A work prepared by a U.S. Government officer or employee as part of his or her official duties is not eligible for U.S. copyright; however foreign copyright laws may differ. Thus this form should be signed even by U.S. Government officers or employees. Signing of this form will not affect U.S. Copyright law provisions in the case of works of the U.S. Government. **If all the authors are in this category, check the box here and return the signed form.** ☐

See page 2 (over) for rights granted by AVS to authors.

2/96

INDEX TO ADVERTISERS

Adhesion International, Quad Group Co.	Cov 4
Ametek-Process & Analytical Instruments Div	A11
International Radiation Detectors Inc.	A4
Leybold Inficon Inc.	Cover 2
Leybold Vacuum Products Inc	Cover 3
LK Technologies	A9
MKS Instruments	A1
Omicron	A11
Stanford Research Systems	A7
Thermionics Northwest, Inc.	A9
Vat, Inc.	A3

Advertising Sales Office

American Institute of Physics
500 Sunnyside Boulevard
Woodbury, NY 11797-2999
Telephone (516) 576-2440
Fax (516) 576-2481

Advertising Manager:

Advertising Sales Representative:

Production Manager:

Richard T. Kobel

Robert G. Finnegan

Betty Aroesty